12-2016

Statistical modeling and simulation of variability and reliability of CMOS technology

Khaled Hassan MD
Purdue University

Graduate School Form
30 Updated 12/26/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Md Khaled Hassan

Entitled

STATISTICAL MODELING AND SIMULATION OF VARIABILITY AND RELIABILITY OF CMOS TECHNOLOGY

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Kaushik Roy
Chair

Anand Raghunathan

Byunghoo Jung

Vijay Raghunathan

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Kaushik Roy

Approved by: Venkataramanan Balakrishnan 12/2/2016

Head of the Departmental Graduate Program Date

STATISTICAL MODELING AND SIMULATION OF VARIABILITY AND RELIABILITY OF CMOS TECHNOLOGY

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Md Khaled Hassan

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2016

Purdue University

West Lafayette, Indiana

Dedicated to my parents, Mohammad Nazmul Haque and Khaleda Akter, who dedicated their whole lives to the well-being of their children

ACKNOWLEDGMENTS

I would like to take this opportunity to express my sincere gratitude to everyone who made this thesis possible.

First and foremost, I thank my advisor Professor Kaushik Roy for being my guide and mentor for the past four years. His valuable advice and support helped me to reach this milestone in my research. His insight and expertise in my research area, and consistent feedback through out this period, helped me to develop technical, problem-solving, and professional skills that will serve me well in my post Ph.D. life. His "MOS VLSI Design" and "Advanced VLSI Design" courses helped me to develop a foundation and insight into VLSI devices and circuit analysis.

I gratefully acknowledge the valuable and constructive feedback from Professor Ashraful Alam on various topics in the field of Reliability Physics in nano electronics. His expertise and exemplary work in this research area motivated me to carry out my own research. His course "Reliability Physics of Nanotransistors" is a great source of knowledge for anyone interested in pursuing research in Reliability Physics.

I would like to thank Professors Anand Raghunathan and Vijay Raghunathan for their valuable suggestions and Dr. Chih-Hsiang Ho and Dr. Sooyoun Kim for being my mentors during my early years as a Ph.D. student. I am also thankful to Dr. Muhammad Masuduzzaman for his guidance and assistance in various advanced topics in reliability physics.

I would like to thank Dr. Arun Goud, Ahmed Kamal Reza, Dr. Selin Mungan, Zubair Al Azim, and Syed Shakib Sarwar for their support and help on various occasions.

Next, I want to thank my parents, Mohammad Nazmul Haque and Khaleda Akter for their selfless love and for providing me with the best possible education with their limited resources. I want to thank my brother Mohammad Sajjad Hossain for his

support, love, and suggestions as an exemplary figure in my life. In addition, I would like to thank all my teachers, relatives, colleagues, and friends. Without their continuous support, I would not have made it this far.

Finally, I would like to thank Professors Anand Raghunathan, Byunghoo Jung, and Vijay Raghunathan for serving on my advisory committee.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

SYMBOLS

ΔV_{th}	Threshold Voltage Degradation
κ	Dielectric Constant
m^*	Effective Mass
γ	Voltage Acceleration Factor
σ	Standard Deviation
β	Weibull Slope
P_F	Failure Probability
μ	Average or mean value
Φ_{Be}	Barrier height of the electrons
T_{hfox}	Thickness of High-κ layer
T_{sox}	Thickness of Interfacial SiO_2 layer
T_{di}	Thickness of a dielectric layer
E_a	Activation Energy
E_r	Relaxation Energy
α	Time exponent of trap generation rate in an oxide layer
a_0	Defect size in an oxide layer

ABBREVIATIONS

VLSI	Very-Large-Scale Integration
CMOS	Complementary Metal-Oxide-Semiconductor
HKMG	High-κ Metal Gate
BTI	Bias Temperature Instability
HCI	Hot Carrier Induced Degradation
RDF	Random Dopant Fluctuation
RTN	Random Telegraph Noise
TDDB	Time Dependent Dielectric Breakdown
SILC	Stress Induced Leakage Current
CDF	Cumulative Distribution Function
TCAD	Technology Computer Aided Design
MC	Monte Carlo
PTM	Predictive Technology Node
LUT	Look-Up Table
WD	Weibull Distribution
GLD	Gate Length Deviation
LER	Line Edge Roughness
DT	Direct Tunneling
WKB	Wentzel-Kramers-Brillouin
TAT	Trap-Assisted Tunneling
ROSC	Ring Oscillator
SRAM	Static Random Access Memory

ABSTRACT

Hassan, Md Khaled Ph.D., Purdue University, December 2016. Statistical Modeling and Simulation of Variability and Reliability of CMOS Technology. Major Professor: Kaushik Roy.

The introduction of High-κ Metal Gate transistors led to higher integration density, low leakage current, and faster switching speed. However, this transition in technology roadmap brought about new failure mechanisms such as Positive Bias Temperature Instability and Stress Induced Leakage Current. In addition, the relentless downscaling of devices to keep up with Moore′s law has significantly increased the time-zero variability caused by Random Dopant Fluctuation, Mean Gate Length Deviation, and Line Edge Roughness. Because of their possible correlation with time dependent aging effects, the quantification of reliability has become more complex than ever. To that effect, we propose a framework to analyze the stochastic nature of different reliability issues and their impact at the circuit level. Our Bias Temperature Instability model for trap generation in NMOS transistors captures both time dependent threshold voltage degradation as well as dielectric breakdown of oxide layers. In addition, we incorporated our model into a Technology Computer Aided Design device simulator in order to capture the correlation between time-zero and time-dependent variability. We observed that the correlation is small, yet significant in the device lifetime prediction. In order to accurately estimate the total guard band necessary for a variation aware Integrated Circuit design, process induced time-zero variation and different aging effects need to be considered simultaneously. In addition, the FinFET architecture can cause additional degradation of the transistors. The field-crowding and 2D-hydrogen diffusion at the fin corners can make the reliability assessment more complicated compared to the planar counterpart. We also proposed a unified model

and simulation framework for reliability analysis in tri-gate FinFET devices. The proposed models and simulation framework can be equally applicable to sub-10nm technology nodes.

1. INTRODUCTION

Fast and reliable! Apart from the cost, these are the only two aspects we tend to think of when we purchase an electronic gadget or laptop: how fast it can run and how long will it provide functionality without losing performance? Over the last decade, we have witnessed unprecedented development in Very-Large-Scale Integration (VLSI) technology that makes our lives easier and faster than ever. One or two devices can take care of most our daily needs. Almost all of us have observed the gradual degradation in the performance of electronic devices over time. In other words, we see a time dependent loss of performance or even a complete failure of these electronic devices. In order to predict the lifetime of a laptop with billions of transistors, it is necessary to have very robust theories that can analyze various reliability aspects with incredible accuracy. Thanks to the researchers who have been working in the field of reliability physics for many years, we have achieved quite a few milestones to date. However, as consumers, we always want faster and more reliable products. Reliability always comes at the expense of speed and vice versa. Hence, manufacturers have to reach an optimum condition where they can meet consumer expectations of performance with an acceptable lifetime.

Due to high integration density, large off-state leakage current, and high on-die temperature, the issues of variability and reliability are of the utmost importance in todays technology and expected to only become more crucial in the future. Hence, the design of a variation-tolerant system requires an expert understanding of device physics, the nature of different variability and reliability issues, and any possible correlations among them. In scaled technologies, the time-zero variability happens mostly due to the phenomenon called Random Dopant Fluctuation (RDF). The small numbers of discrete dopants in the tiny channel regions cause the nominal threshold voltage (V_{th0}) to fluctuate among microscopically similar devices due to the randomness

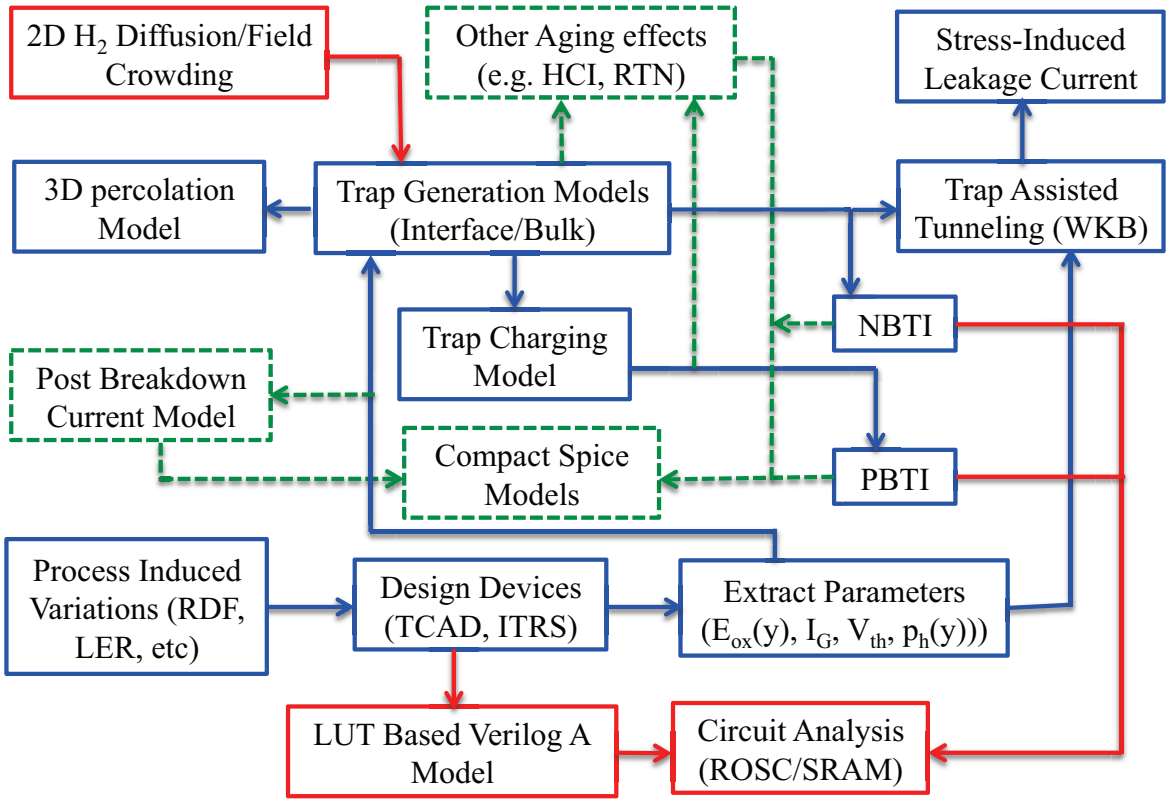


Fig. 1.1. The proposed simulation framework. The dotted lines indicate future extension.

in their number and locations. Apart from RDF, Line Edge Roughness (LER), Line Width Roughness (LWR), Mean Gate Length Deviation (GLD), Work Function Fluctuation (WFF), and Non-Uniform Temperature are some the notable process-induced variations in todays CMOS technology [1–5]. The time dependent, variability-induced failures can be either parametric, which is mostly V_{th} degradation, or catastrophic failure due to the gradual wear of thin oxide layers over time. The V_{th} degradation is attributed to Bias Temperature Instability (BTI) caused by the breakage of interface or bulk oxide bonds due to high operating voltage or temperature [3, 6–14]. Oxide breakdown, also known as time-dependent dielectric breakdown (TDDB) occurs due to the formation of conduction paths through the bulk and interface traps caused by BTI. In addition, Hot Carrier Induced Degradation (HCI), Random Telegraph Noise

(RTN), and Stress Induced Leakage Current (SILC) also contribute to V_{th} and oxide quality degradation [5]. Some of the process induced variability and aging effects in today's CMOS technology are listed in Table 1.1. Both time-zero and time dependent variability are correlated which needs to be properly addressed. The objective of this work is to develop a simulation methodology that incorporates the combined effect of different process induced and time dependent aging into a unified aging model that can be leveraged to generate compact models for circuit simulations. Fig. 1.1 shows the organization of this thesis. We have designed MOSFET devices using a Technology Computer-Aided Design (TCAD) simulator [15] and benchmarked their characteristics with ITRS projection [7]. In our thesis, we have focused on 30nm planar and 14nm tripple-gate FinFET devices. The core of this work is the unified trap generation model that includes both interfaces (Si/dielectric or dielectric/dielectric) and bulk oxides. Positive BTI (PBTI) is a phenomenon that depends on trap generation at the dielectric/dielectric interface or bulk oxide; Negative BTI (NBTI) is mostly related to the Si/dielectric interface traps. The Hot-Carrier Induced degradation depends on both interface and bulk trap generations. The traps generated due to NBTI, PBTI, and HCI contribute to TDDB and SILC [5]. Therefore, the unified model can be utilized for the analysis of all of these aging effects. In addition, the impact of pre-existing defects can also incorporated into the models. The dotted lines in Fig. 1.1 refer to the future extension of this work.

1.1 Contributions and Organization of this Thesis

The key contributions and the proposed future extension of this work is summarized below:

- Proposed a 3D statistical model for PBTI analysis. The model is equally applicable for planar and FinFET devices.
- Leveraged the PBTI model to include TDDB and SILC. For TDDB and SILC analysis, we have implemented the 3D percolation model and $Wentzel-Kramers-$

$Brillouin(WKB)$ [16] approximation-based Trap Assisted Tunneling (TAT) model, respectively.

- Proposed a TCAD-based simulation methodology that accounts for the combined effects of process-induced variation and aging. As an example, we have investigated the impact of the combined effects of RDF and BTI (both positive and negative), TDDB, and SILC at the device level. We have implemented a statistical Reaction-Diffusion (R-D) model for NBTI analysis. The results show that the effects of RDF on BTI need to be considered for estimation of performance degradation. However, the impact on TDDB and SILC analysis is almost negligible.

- From the TCAD simulator, we have generated a Look-Up Table (LUT) based Verilog A model [17] for Spice simulation and investigated the impact of RDF on BTI in circuit performance. In our analysis, we have focused on the oscillation frequency degradation of Ring Oscillators (ROSC) and different failure mechanisms of 6-Transistors Static Random Access Memory (6T-SRAM). We observe that the effect of RDF on BTI decreases with the increase in the number of ROSC stages. However, this correlation needs to be considered in SRAM failure analysis in order to achieve high error immunity.

- Leveraged the simulation framework developed for planar devices to FinFET technology, investigated the impact of the fin shapes on device performance and oxide electric field, and extended the RD model to incorporate the effect of 2D H_2 diffusion and fin thickness in NBTI analysis. We have implemented a unified trap generation model for PBTI, NBTI, and TDDB and verified the model with experimental results.

In addition to this introductory chapter, there are five other chapters in this thesis. In Chapter 2, we propose our stochastic BTI model for high-κ NMOS transistors. The model takes into consideration each trap location and their distinct impact on V_{th} degradation. The model is flexible and equally applicable for advanced technology

Table 1.1.
Variabilities in nano-scale transistors

	Name	Short Description
Time-Zero or Process Induced Variability	Random Dopant Fluctuation (RDF)	Due to the randomness of number and location of discrete dopants in the channel region
	Line Edge/Width Roughness (LER/LWR)	Molecular structure of the photoresist causes deviation from straight lines
	Gate Oxide Thickness, Channel Length, Work Function Fluctuations	Due ot the deviations of oxide thickness, channel length, and gate work function from the nominal value
Time Dependent or Aging Induced Variability	Positive/ Negative Bias Temperature Instability (P/N BTI)	Due to the generation and/or charging/discharging of defects at the Si/dielectric interface, dielectric/dielectric interface, and/or in the bulk oxide
	Hot Carrier-Induced damage (HCI)	Defects generation by the high energy carriers due to the application of high drain bias
	Stress-Induced Leakage Current (SILC)	Increase of power consumption due to Trap-Assisted Tunneling current
	Random Telegraph Noise (RTN)	Random trapping/de-trapping of carriers in defects
	Time Dependent Dielectric Breakdown (TDDB)	Formation of continuous gate-body leakage current paths due to the generation of defects
	Radiation-Induced Damage (RID)	Damage induced to the transistors by radiation (e.g., alpha particles, cosmic ray, etc.)

(such as FinFETs). In Chapter 3, we propose a simulation methodology in order to analyze the impact of process induced time-zero variation and different time dependent aging effects. We have integrated our PBTI model with a TCAD simulator. In addition, we have extended the PBTI model to incorporate SILC and TDDB analysis within the same simulation framework. We observe that the impact is somewhat significant on BTI variability. However, TDDB and SILC analysis show negligible dependance on RDF. We have generated LUT-based Verilog A model from the TCAD simulations in order to carry out the circuit analysis. We observe from our circuit simulations that the impact of correlation between RDF and BTI is significant when the mean BTI degradation is small (50 mV or less). The correlation becomes weaker with the increase of circuit complexity. In Chapter 4, we carry out SRAM failure analysis using the simulation mythology proposed in Chapter 3. The results show that the correlation needs to be considered in order to achieve 99.7% (3σ) or higher confidence level. In Chapter 5, we extend our models to come up with a unified model for reliability analysis of FinFETs. We have investigated the field-crowding effect at the corners of the tri-gate structure and extended the statistical R-D model for NBTI analysis. We have verified the trap generation models for NBTI, PBTI, and TDDB analysis with experimental results. In Chapter 6, we conclude this thesis and discuss some of the future works that can amend our unified model presented in Chapter 5.

2. STOCHASTIC MODELING OF POSITIVE BIAS TEMPERATURE INSTABILITY IN HIGH-κ METAL GATE NMOSFETS

Positive Bias Temperature Instability (PBTI) has become one of the major reliability concerns in the present day CMOS technology. PBTI mostly degrades the performance of high-κ metal gate (HK/MG) nMOSFETs and is dominated by time dependent stress induced trap generation. The PBTI models proposed so far in different literature are deterministic in nature and overlook the randomness of the temporal degradation of transistor threshold voltage, V_{th}. In this chapter, we present a stochastic model for PBTI in order to predict the behavior of HK/MG nMOSFETs under inversion mode stress. The model is scalable, and is verified by experimental data from two different groups. Our model separately considers each trap to predict their impact on device performance. The simulations are carried out at accelerated voltage and temperature conditions and we noted the variations in V_{th}. In addition, we have analyzed the impact of PBTI on the performance of a ring oscillator and concluded that the circuit speed suffers significant degradation due to this effect. We have also observed that the adverse effect of PBTI becomes more severe as we go deeper into the nanometer technology.

2.1 Introduction

With the introduction of high dielectric constant (high-κ) MOS transistors combined with metal gate electrodes, the gate leakage current and the corresponding stand by power dissipation are significantly reduced in comparison to the conventional silicon oxynitride (SiON) transistors [6, 18, 19]. However, the transition from oxynitride/poly-Si gates to high-κ/metal gates, has led to a new failure mechanism,

positive bias temperature instability (PBTI) in n-channel devices [7]. Since HfO_2 bulk traps cannot be charged under negative stress voltage, these defects only affect the threshold voltage and leakage current of NMOSFETs [8]. Although the effect of PBTI was smaller in older technology nodes (65nm or earlier), it is regarded as a major reliability issue in case of HK/MG transistors [20]. The presence of d-shell states, relatively higher coordination number, and low processing temperature of Hafnium (Hf) based oxides lead to drastically different electronic structure than the conventional SiO_2 and make them vulnerable to defect formation [21, 22]. Over the last decade, researchers have argued whether the already existing traps or the stress induced generated defects are responsible for the transistor V_{th} shift due to PBTI effect. The degradation used to be attributed to the charging and discharging of pre-existing traps in earlier literature [13, 22–25]. However, recent research confirms that PBT instability occurs mostly because of the stress induced defect generation and the effect of preexisting traps are only dominant for a short period of time at the beginning of the stress duration [8, 10, 14, 26–28]. A typical HK/MG gate stack (Fig. 2.1) is comprised of a high-κ dielectric layer (such as HfO_2) sandwiched in between an inter-

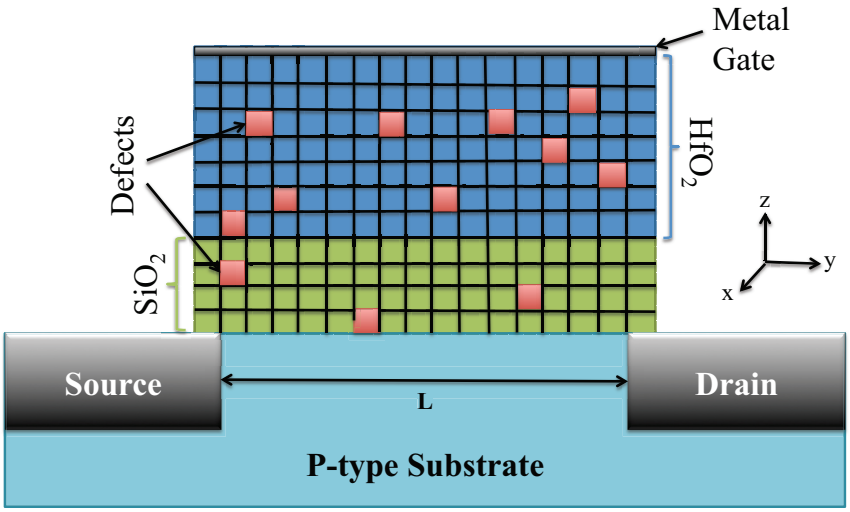


Fig. 2.1. Dual layer HK MG nMOSFET. The HK layer is much more prone to defect formation than the SiO_2 layer.

layer (typically SiO_2) and a metal gate (usually TiN). The interplay of these different elements has significant impact on the reliability assessment of these devices [29]. The defects in the SiO_2 layer have relatively negligible effect on PBT instability and hence, can be ignored for such device structures [14].

Due to continuous downscaling of devices, the minimum feature size has gone down to about a hundred atomic sizes. This increased both the temporal and the process related device parameter variability. As a result, reliability projection has become more difficult. The PBTI models proposed so far in different literature are based on the assumption that the charging probability of each trap is uniform. In reality, they are not uniform. Consequently, they cannot accurately predict the variation in device performance as well as lifetime due to the PBTI effect. A statistical model is therefore needed to precisely predict device lifetime.

2.2 Modeling

The flowchart in Fig. 2.2 shows the framework of the proposed model. The model takes different physical properties (such as device dimensions, oxide properties, etc) and stress conditions as input. In the second step, it generates the temporal trap distribution and determines their charging probability based on the probability theory. In the final step, it determines the transistor threshold voltage degradation depending on the locations of each trap. We have made the following assumptions in our work:

- The traps are generated uniformly throughout the dielectric layer.
- The charging of the traps is non-uniform. In other words, the probability for a trap being occupied depends on the location of that trap.
- We have ignored the lattice mismatch of HfO_2/SiO_2 interface. Lattice constant of SiO_2 and HfO_2 are 0.734nm and 0.498nm respectively. Hence, band alignment can be an issue for both HfO_2/SiO_2 and HfO_2/Metal interfaces [30]

- We did not consider mobility degradation due to the tunneling of carrier from channel to the bulk region.

For our modeling, we have used a hafnium based high-κ dielectric on a SiO_2 interlayer. Unless otherwise specified, W=L=22nm and the thicknesses of the oxide layers are 2.5nm and 1nm respectively, for the simulation results. The trap size in HK layer is

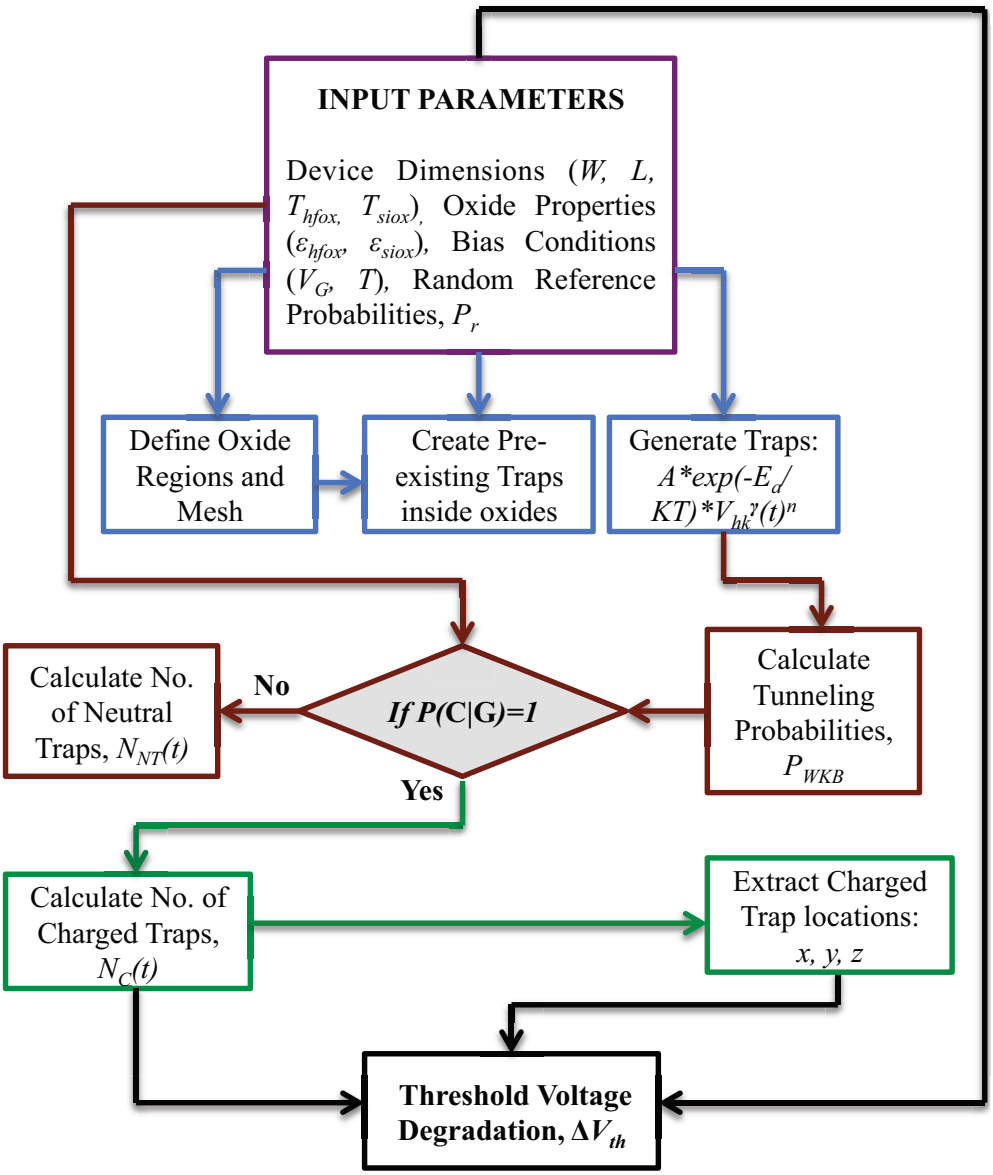


Fig. 2.2. The flowchart of the proposed model.

larger than that in SiO_2, which can be attributed to the difference in energy levels of the traps associated with the defects. The trap level in HfOx is <1.4eV, which is much shallower than that in SiON [21, 25, 31, 32].

2.2.1 Trap Generation

In our proposed model, trap generation process considers the impact of the following factors:

- Stress time
- Voltage acceleration
- Temperature acceleration

Depending on the results of accelerated test conditions, the device performance at operating condition can be projected. The number of traps generated due to PBTI

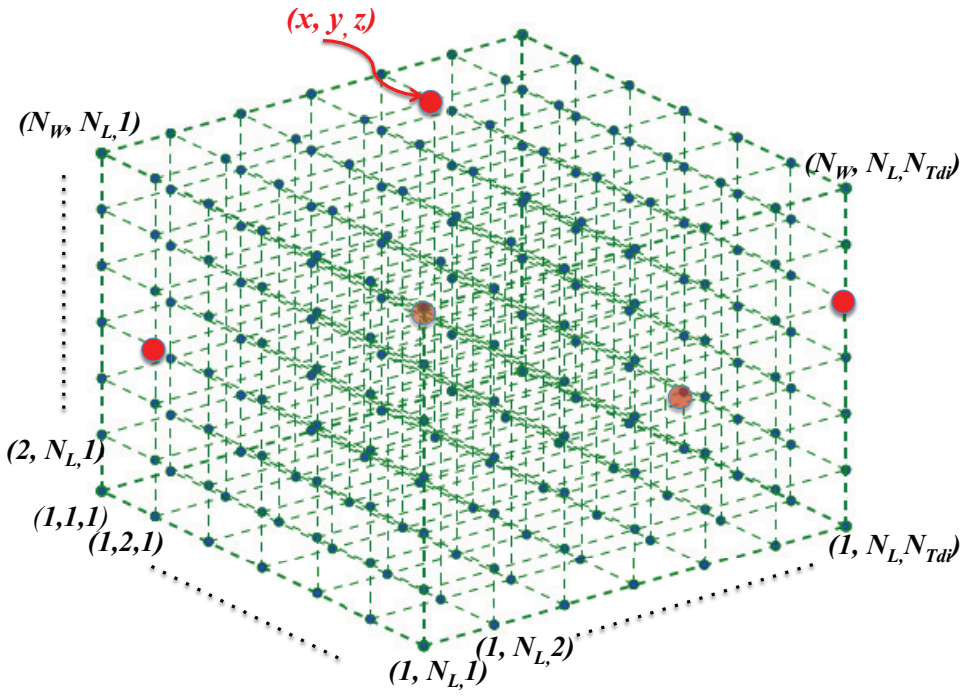


Fig. 2.3. A 3D Grid representing a sample oxide layer (defect locations are shown in red).

affects the transistor threshold voltage. The change in threshold voltage can be expressed as follows:

$$\Delta V_{th}(V_G, T, t) = f(N_C(V_G, T, t)) = f(f(N_G(V_G, T, t))) \tag{2.1}$$

where, N_C and N_G are the number of charged and generated traps, respectively, for a certain bias condition (Gate Voltage: V_G, Temperature: T) at a specific time, t. Based on the ΔV_{th} model in [8,33], the number of traps generated in the i_{th} dielectric layer, N_{Gi} of a stack can be modeled by a power law:

$$N_{Gi}(V_G, t) \propto (V_{Gi})^{\gamma} t^n \tag{2.2}$$

where V_{Gi} is the voltage drop across the same dielectric layer, γ is the voltage acceleration factor, and n is the PBTI power exponent. The typical value of γ is in the range of 5 to 10 [8, 14] while n is in between 0.1 and 0.2 [8]. The defect generation is strongly temperature dependent [8, 27] and can have a significant role in device performance [10, 14]. Hence, the temperature dependence can be modeled by the Arrhenius equation [34]:

$$N_{Gi}(T, t) \propto exp(-E_a/K_B T) t^n \tag{2.3}$$

where K_B is the Boltzmann constant and E_a is the activation energy (usually 1-1.5 eV [10, 14, 26]) for the breakage of oxide bonds in HfO_2 dielectric layer. We can combine (2.2) and (2.3) to model the number of generated traps N_{Gi} as follows:

$$N_{Gi}(V_G, T, t) = A(V_{Gi})^{\gamma} exp(-E_a/K_B T) t^n \tag{2.4}$$

where A is a calibration parameter that depends on both technology node and device structure. For modern high-κ metal gate transistors, i indicates either SiO_2 IL or a high-κ dielectric layer (such as HfO_2). We assumed a random distribution of these traps across the HfO_2 layer. A pseudo-stochastic process $S_j(t)$ is assigned to each of the oxide bonds. $S_j(t)$ can be either 0 or 1 if the bond is tied or broken, respectively. Fig. 2.3 shows a 3D grid structure for the simulation. N_w, N_L, and N_{Tdi} are the

number of points taken along the width, length, and thickness of the i_{th} dielectric layer, respectively. For our simulation, we have assumed the defect size of HfO_2 as 0.4-0.5nm. The temporal variation of the randomly distributed generated traps therefore can be represented by the following equation:

$$N_{Gi}(t) = \int_{x=0}^{W} \int_{y=0}^{L} \int_{z=0}^{T_{di}} S_{xyz}(t) \mathrm{d}x \mathrm{d}y \mathrm{d}z \tag{2.5}$$

where, W, L, and T_{di} are respectively the width, length, and, the thickness of the oxide layer.

2.2.2 Trap Charging

The band diagram in Fig. 2.4 shows the operation of n-MOSFET in the inversion mode. In this diagram, V_{sox} and V_{hfox} indicate the voltage drop across the SiO_2 and HfO_2 layers, respectively. ϕ_B and ϕ_{hfox} denote different barrier heights seen by the tunneling electrons and ϕ_T represents the trap energy level inside the high-κ layer. With the generation of new electron traps, the probability of charging enhances and we need to know the exact location of these traps in order to determine the trap occupation effect of each of them separately. Trapping of a carrier is a conditional event and is dependent on the generation of traps. If P_G and P_C represent the probability of trap generation and trap charging respectively, the probability that both events occur at the same location (x, y, z) at a specified time is given by [35]:

$$P(C|G) = P(C \bullet G)/P(G) \tag{2.6}$$

where G and C are trap generation and charging events respectively and $P(C \bullet G)$ is the probability of G and C taking place at the same time. $P(C|G)$ in (2.6) is the probability of charging a trap at (x, y, z), given a trap is being generated at the same location. Since $C \subset G$, (2.6) can be simplified to

$$P(C|G) = P(C)/P(G) \tag{2.7}$$

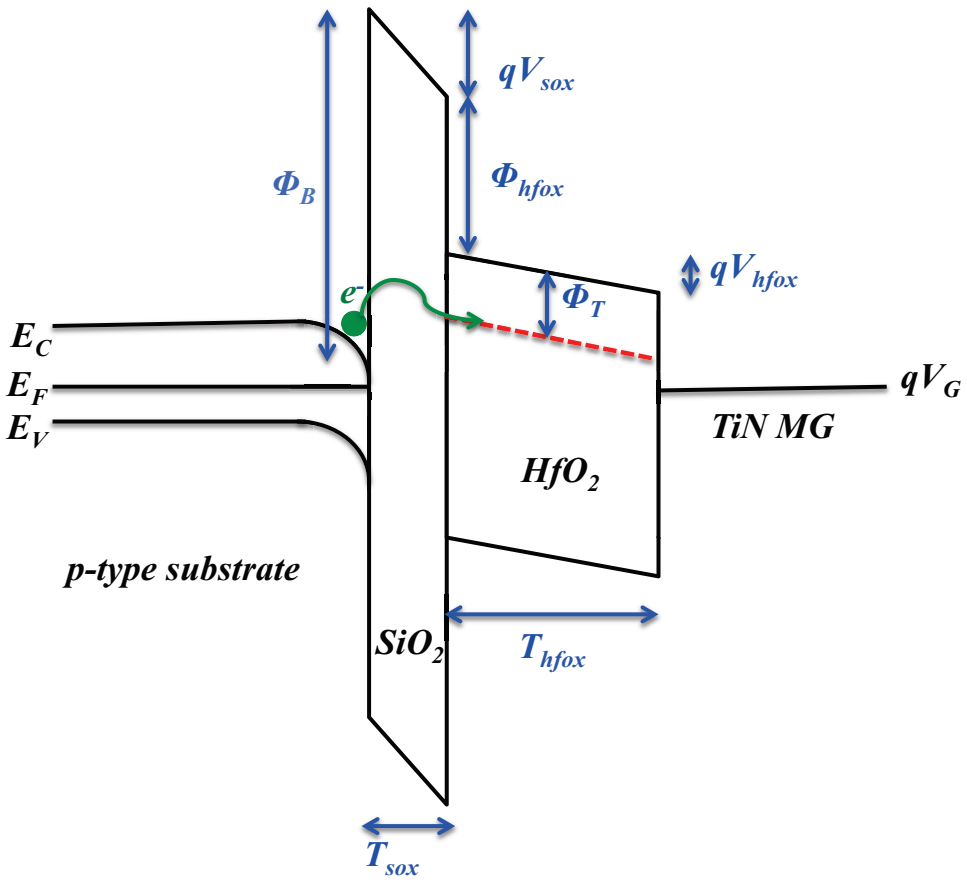


Fig. 2.4. Band diagram of nMOSFET in inversion mode. The numerical values of different parameters are given in Table 2.1.

The boundaries of (2.6) and (2.7) is determined by the defect size and lattice structure and hence, by the maximum number of generated oxide traps. The number $N_W.N_L.N_{Th}$ (Fig. 2.3) limits the maximum number of defects generated for a given device. In order to model the conditional probability $P(C|G)$, we calculated the tunneling probability of electrons, P_{WKB} (considering substrate injection only) by the $Wentzel - Kramers - Brillouin(WKB)$ method [16, 37, 38] and compared that with a set of pseudo-random probabilities P_r.

$$P_{WKB} = exp\Big(\frac{-2}{\hbar}\Big[\int_0^z \sqrt{2m^*_{de}(\phi_{Be} - E_k)}\mathrm{d}z\Big]\Big) \tag{2.8}$$

Table 2.1.
Different parameters used in our PBTI simulations

Parameters	Value	Reference	Parameters	Value	Reference
N_A/cm^3	2.7x10^{18}	[36]	E^C_{hfox}	1.4eV	[21]
n	0.1-0.2	[11, 12, 14]	ϕ_T	<1.4eV	[32]
γ	5-10	[11, 12, 14]	ϕ_B	3.1eV	[16,36]
m^*_{sox}	$0.5m_0$	[37]	E_a	0.11-0.15	[10,26]
m^*_{hfox}	$0.18m_0$	[36]	V_F	-1V	[36]
N_A : Doping density in Si substrate m_0 : The mass of electron in free space E^C_{hfox} : Conduction band offset of HfO_2 on Si					

where, m^*_{de}, ϕ_B, $\hbar$, E_k are the effective mass of electron in dielectric layers, effective barrier height seen by the tunneling electron, reduced Plancks constant, and the average kinetic energy of electron in 3D semiconductor, respectively. The voltage drop across the dielectric stack $V_{stk}(= V_{sox} + V_{hfox})$ is calculated as follows [37,38]:

$$V_{stk} = V_G - V_F - \phi_S \tag{2.9}$$

Where, V_F is the flat band voltage and ϕ_S is the surface potential for the device. If we define $C_r = C_{sox}/C_{hfox}$, the voltage drop across the SiO_2 and HfO_2 layers can be determined as:

$$\begin{aligned} V_{sox} &= \frac{V_{stk}}{C_r + 1} \\ V_{hfox} &= \frac{V_{stk}}{1/C_r + 1} \end{aligned} \tag{2.10}$$

where C_{sox} and C_{hfox} are the per unit capacitances of SiO_2 and HfO_2 layers, respectively. Once we have P_r and P_{WKB}, $P(C|G)$ can be modeled as follows:

$$P(C|G) = \begin{cases} 1 & \text{if } P_{WKB} > P_r \\ 0 & \text{Otherwise} \end{cases}$$

Since we can extract the generated trap locations from our simulation, $P(G)$ is either 0 or 1 if a bond is tied or broken, respectively. In addition, $P(C) = 0$ if the bond is tied. Therefore, at a certain time t, (2.7) becomes:

$$P(C)_t = P(C|G)_t \tag{2.11}$$

We calculate the total number of charged traps using the following expression:

$$N_C(t) = \int_{x=1}^{N_W} \int_{y=1}^{N_L} \int_{z=1}^{N_{T_{di}}} P(C)_{t,x,y,z} \mathrm{d}x\mathrm{d}y\mathrm{d}z \tag{2.12}$$

where $P(C)_{t,x,y,z}$ in (2.12) indicates the charging probability of a trap located at (x, y, z) at a certain time t.

2.2.3 V_{th} Degradation

Our model individually considers each charged defect. For the number of charged traps evaluated by (2.12), we can determine the PBTI induced threshold voltage degradation. The ΔV_{th} in the high-κ layer is determined using *Poisson's equation*:

$$\begin{aligned} {\Delta V_{th}}^{hk} &= \frac{-q}{WL} \int_z \frac{N_C(t)}{C_{hk}(z)} \mathrm{d}z \\ &= \frac{-q.N_C(t)}{(WL).\epsilon_{hk}\epsilon_0} \int_z (T_{hfox} + T_{sox} - z) \mathrm{d}z \end{aligned} \tag{2.13}$$

where ϵ_{hk} is the relative dielectric constant of the HfO_2 layer. From (7) and (11), it can be concluded that the charging event C dominates the total number of charged traps as well as the final threshold voltage degradation in (13).

2.3 Results and discussion

In this section, we validate our model by comparing with experimental and simulation data from two different groups and explain the statistical nature of PBTI degradation. The results show a wide variation of V_{th} degradation. We also note wider variations with the increase in stress voltage and/or temperature.

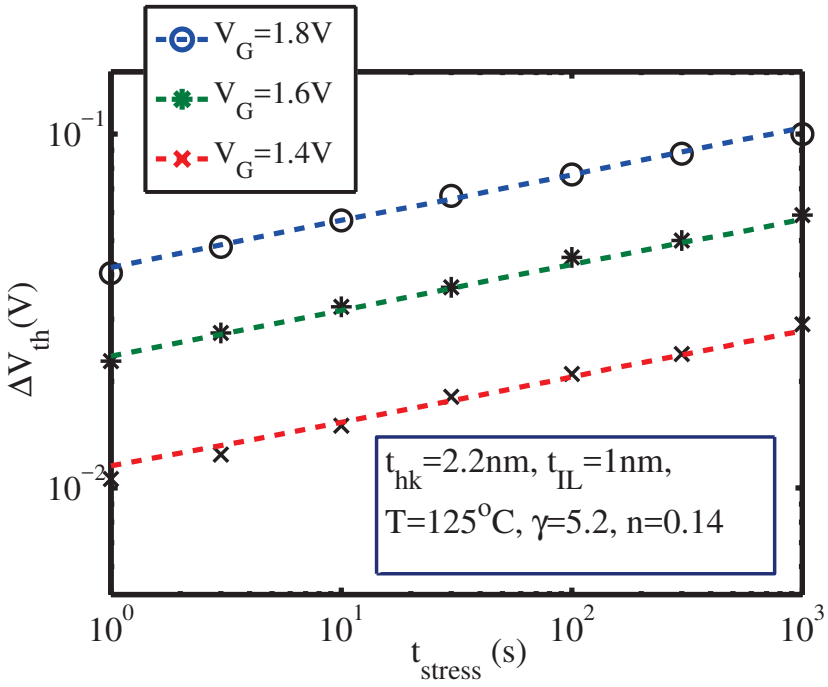
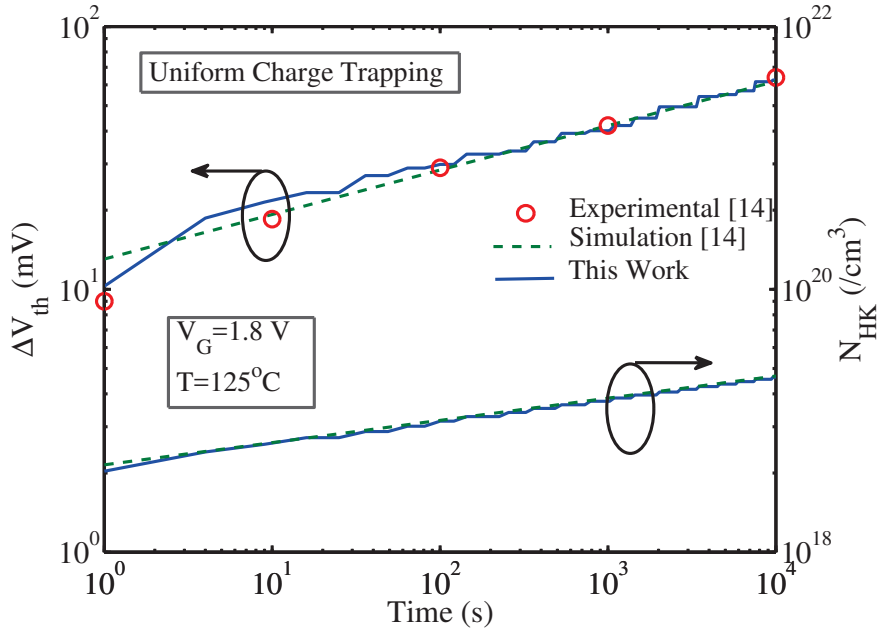


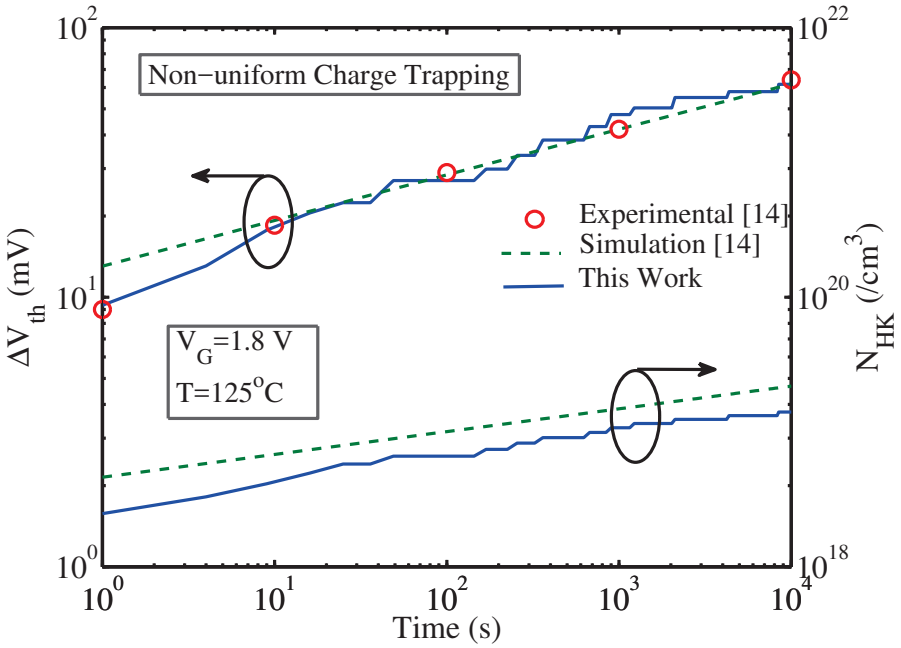
Fig. 2.5. Average ΔV_{th} shift under different stress voltages for 1000 samples (W=L=22nm). Experimental results are taken from [8].

2.3.1 Model Validation

There have been discrepancies among the experimental results reported by different groups and we see a wide range of variation in voltage acceleration factor, PBTI power exponent, and the level of V_{th} degradation among different references [8, 10, 14, 22]. Therefore, it is necessary to come up with a model that can provide adaptability using minimum possible fitting parameters. In our model, only the parameter $'A'$ in (2.4) is calibrated to match the experimental data published by two different groups. We have ignored the carrier hoping among different traps since this process is considered very slow in comparison to relatively fast trapping of the traps [22]. In all cases, we have taken the temperature activation factor to be 0.11 eV [14].

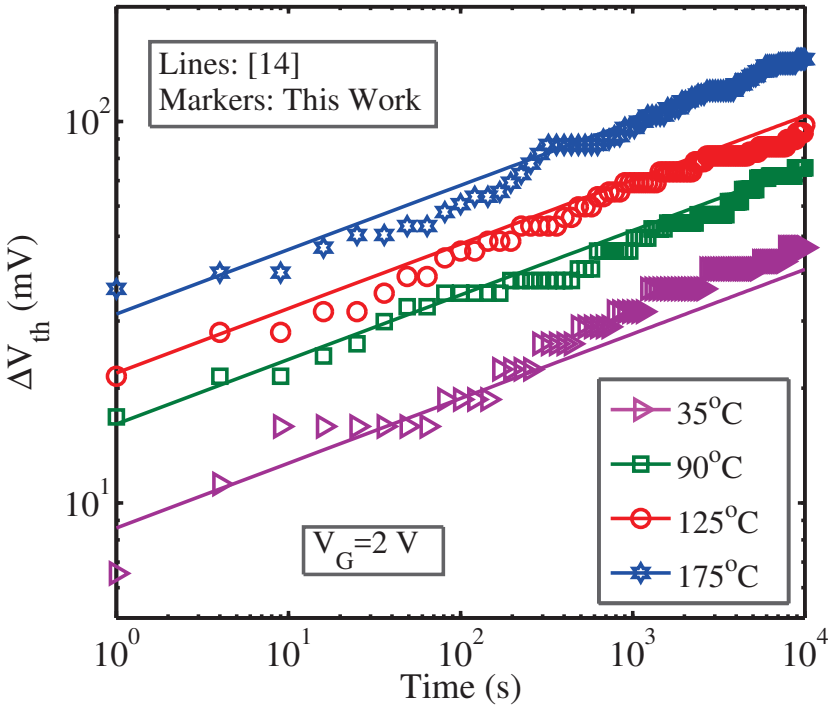


(a)

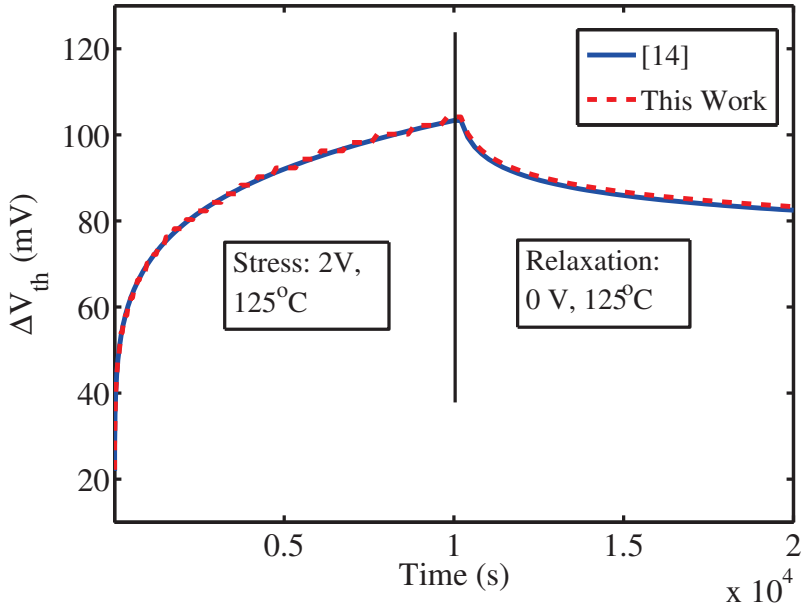


(b)

Fig. 2.6. Comparison between: (a) Uniform; (b) Non-uniform charging probability of traps. The uniform charging shows over estimation of the trap density in side the oxide layer.



(a)



(b)

Fig. 2.7. (a) Variation of ΔV_{th} at different temperatures; (b) The relaxation effect following a 2V and 125^{o}C stress. The results are in good aggreement with [14].

Fig. 2.5 shows the V_{th} degradation for different stress voltages at 125^oC. Our simulation results are in good agreement with A. Kerber, *et al.* [8]. The results reflect strong voltage dependence of trap generation and electron trapping. In Fig. 2.6, the results are compared to Yang *et al.* [14]. The power exponents were taken as 0.17 and 5 for the PBTI power exponent and voltage acceleration factor, respectively. Fig. 2.6(a) and (b) shows how the simulation can lead to erroneous results (wrong number of defects) if we consider uniform charging probability for all traps. It can be clearly seen in Fig. 2.6(b) that the charge density does not match with [14] if we consider non-uniform charging probability for different traps. Reason being the traps close to the channel has higher probability of being occupied and larger contribution to V_{th} degradation than those relatively far away from the interface.

Fig. 2.7 (a) verifies the simulation results for different temperature stress. Since PBTI is a temperature-activated process, the increase in temperature generates larger number of traps according to (2.3). In addition, the kinetic energy of conduction band electrons also increases, which in turn reduces the barrier height seen by the carrier. As a result, both trap generation and trapping mechanisms are accelerated and we see significant enhancement in performance degradation with the increase in temperature. Fig. 2.7(b) shows the recovery effect preceded by a stress condition of 2V and 125^oC. The mean ΔV_{th} of 1000 samples matches very well with the result reported in [14]. The PBTI relaxation effect follows $log(t)$ dependence on time t [33]. Since, BTI is a frequency independent phenomenon [39], the duty cycle dependance of trap generation can be represented by the following expression:

$$N_{OT/IT}(f, s_p) = a(f, s_p) * N_{OT/IT}(0, 1) \tag{2.14}$$

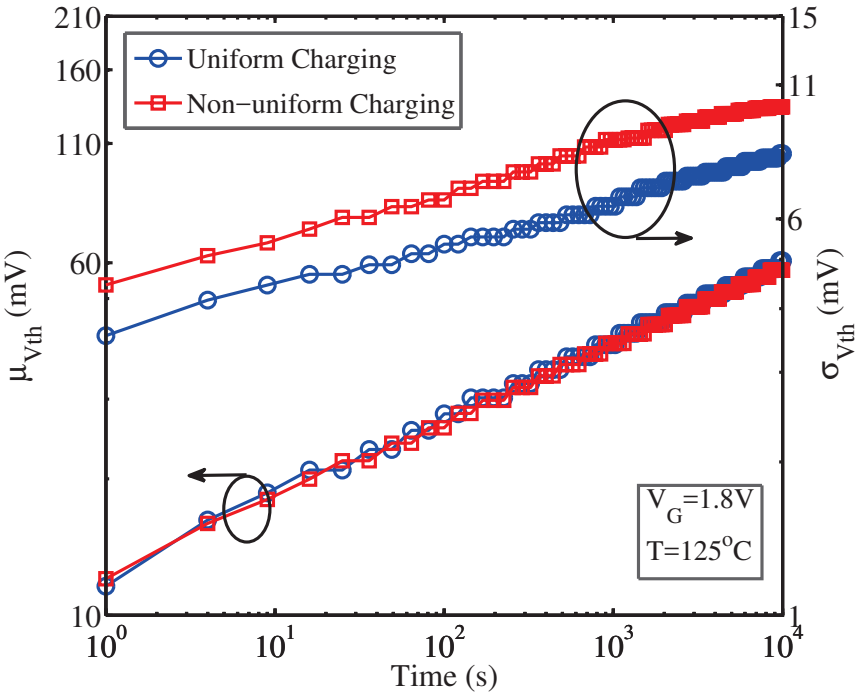
where, the factor $a(f, s_p)$ denotes the f and s_p dependance of BTI effects and $N_{OT/IT}(0, 1)$ is the traps generated due to BTI degradation under DC stress. From Fig. 2.7 (b), the factor $a(f, s_p)$ for 50% duty cycle is extracted as 0.75. NBTI shows similar duty cycle dependance and the corresponding $a(f, s_p)$ factor is 0.796parameter is [40].

2.3.2 Statistical Distribution of V_{th} Degradation

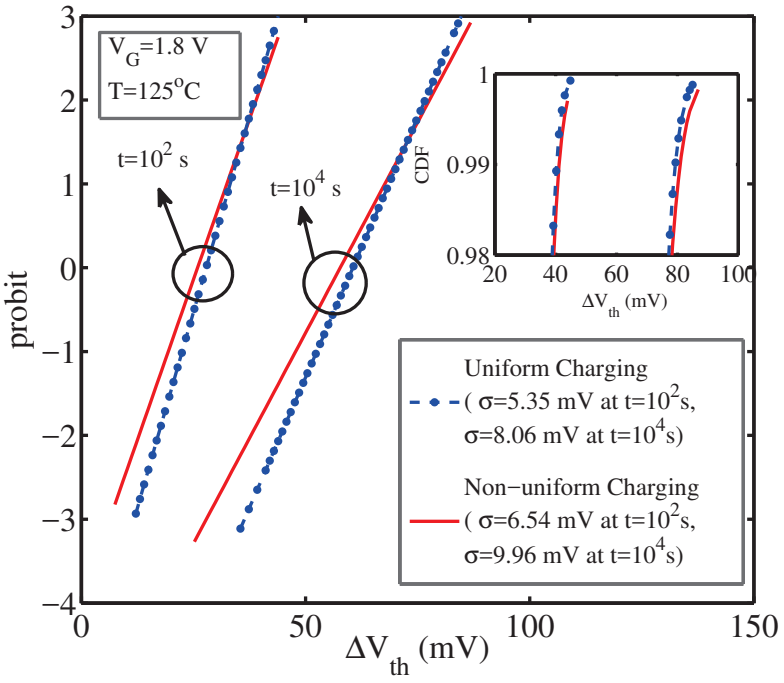
In Fig. 2.8 (a), we plotted $\sigma_{V_{th}}$ and $\mu_{V_{th}}$ for both uniform and non-uniform charging of traps over a period of 10000 seconds for 1000 sample devices and Fig. 2.8 (b) shows the corresponding probit plot and the upper tail of the cumulative distribution function (CDF) at two different stress time ($10^2 s$, and $10^4 s$). The results suggest that the standard deviation in the case of uniform charging of traps is lesser than the non-uniform case. The average value of $\Delta V_{th}(t)$ however, remains almost the same. This leads to an underestimation of the temporal V_{th} degradation for uniform charging. In Fig. 2.9 (a and b), we show the effect of voltage and temperature variation on PBTI. The Monte Carlo (MC) simulation is carried out for 10000 samples. The results indicate significant increase in both $\sigma_{V_{th}}$ and $\mu_{V_{th}}$ at higher stress voltage and temperature. In addition, we have noticed that the power exponents (both voltage acceleration factor and time exponent) remain same for the average $\Delta V_{th}(t)$ ($\gamma_\mu \approx \gamma$ and $n_\mu \approx n$). These power exponents however, are smaller by a factor of 2 for the standard deviation ($\gamma_\sigma \approx \gamma/2$ and $n_\sigma \approx n/2$). This important observation is in agreement with what has been reported for NBTI in [42] and can help us to project the PBTI effect for any stress period for a given stress condition using (2.14):

$$\begin{aligned} \mu(t_2) &= exp(ln(\mu(t_1) + (ln(t_2) - ln(t_1))n_\mu) \\ \sigma(t_2) &= exp(ln(\sigma(t_1) + (ln(t_2) - ln(t_1))n_\sigma) \end{aligned} \tag{2.15}$$

where t_1 and t_2 are the times from and to when the degradation is projected. The effect of technology scaling is shown in Fig. 2.10. In Fig. 2.10(a) we observe an increase in the variation with the decrease in device area. Hence a higher variation is expected as the device size is scaled down. In Fig. 2.10(b), we have plotted the PBTI variability against $\frac{1}{\sqrt{WL}}$. We see a linear relationship which confirms that the variation follows Pelgrom's rule [41].

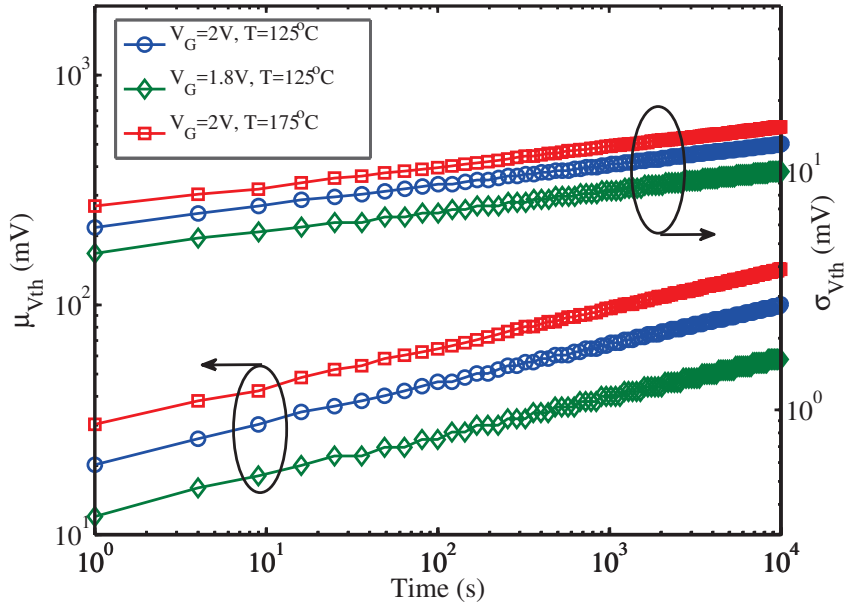


(a)

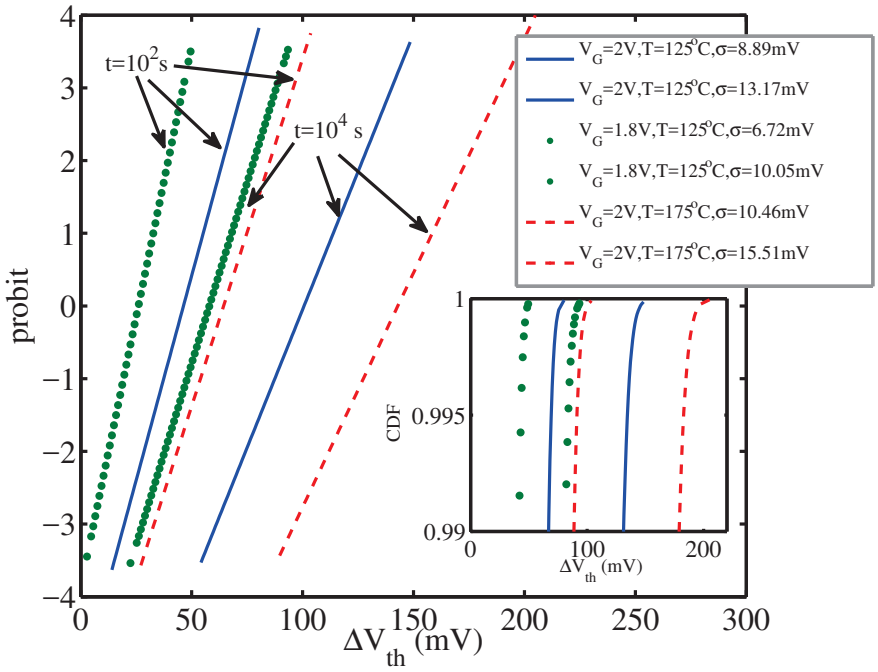


(b)

Fig. 2.8. Statistical Distribution of ΔV_{th} degradation for both uniform and non-uniform charging probability: (a)$\mu_{\Delta V_{th}}$ and $\sigma_{\Delta V_{th}}$; (b)The probit plot and the upper tail of the CDF at $t = 10^2$ and $t = 10^4 s$.

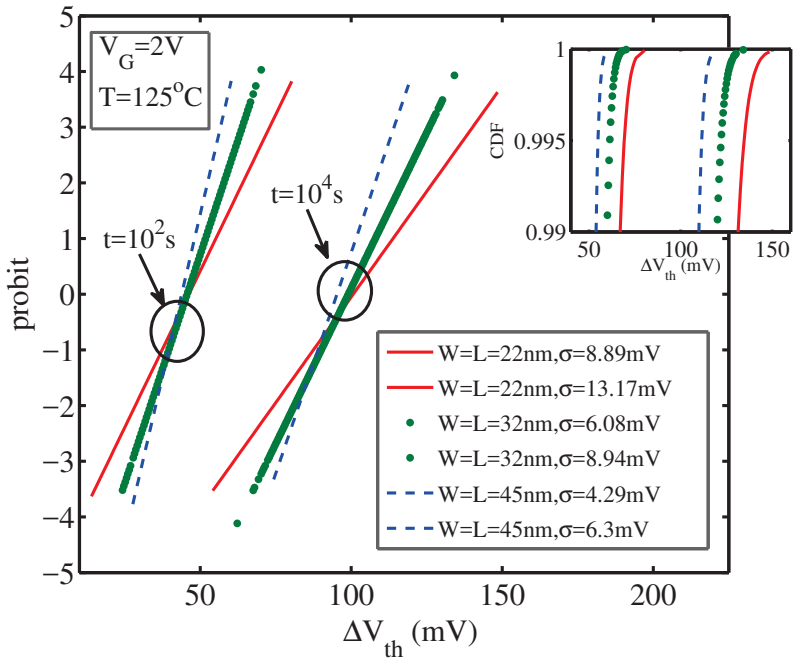


(a)

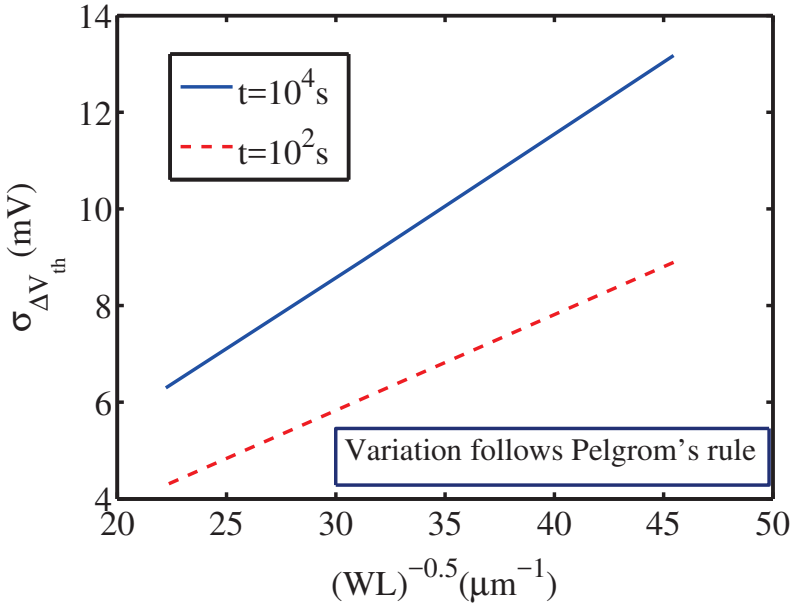


(b)

Fig. 2.9. Statistical Distribution of ΔV_{th} for different voltage and temperature: (a)$\mu_{\Delta V_{th}}$ and $\sigma_{\Delta V_{th}}$; (b)The probit plot and the upper tail of the CDF at $t = 10^2$ and $t = 10^4 s$.



(a)



(b)

Fig. 2.10. ΔV_{th} degradation of 10000 samples for different technology nodes: (a) at $t = 10^2$ and $t = 10^4 s$; (b) Verification of Pelgrom's rule [41]. The temporal variation increases with the decrease in device dimensions.

2.3.3 Comparison with NBTI distribution

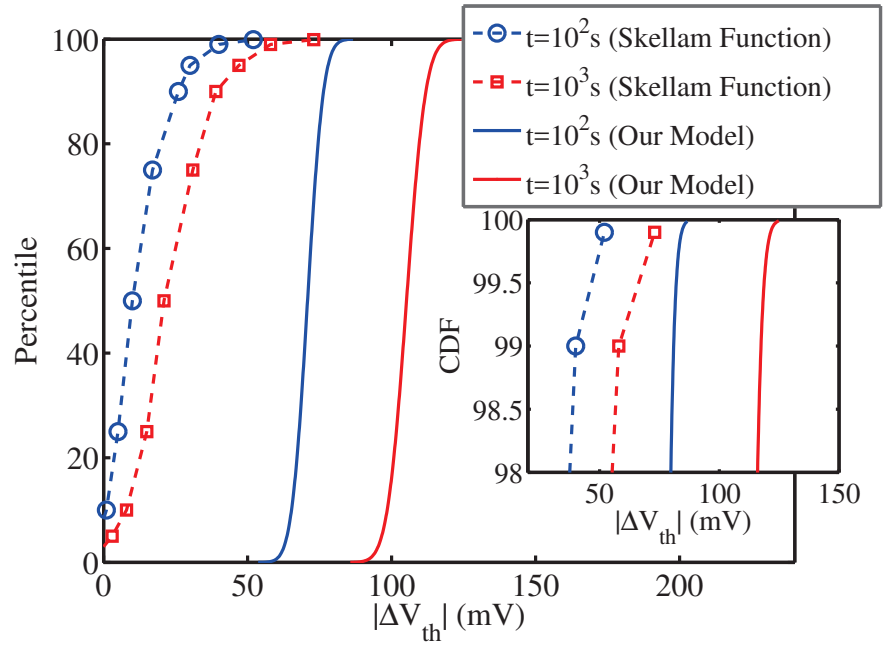
Fig. 2.11 shows the comparison of our PBTI model with two different NBTI models. The insets show the upper CDF tail of the corresponding distributions. The data extracted from [43] (Skellam Function) and [44] (Defect Centric Approach) show that NBTI for p-MOSFET has a wider distribution tail than that of PBTI for high-κ n-MOSFET. The width and effective length for these simulations are 90nm and 35nm, respectively.

2.3.4 The Effect of PBTI on Circuit Performance

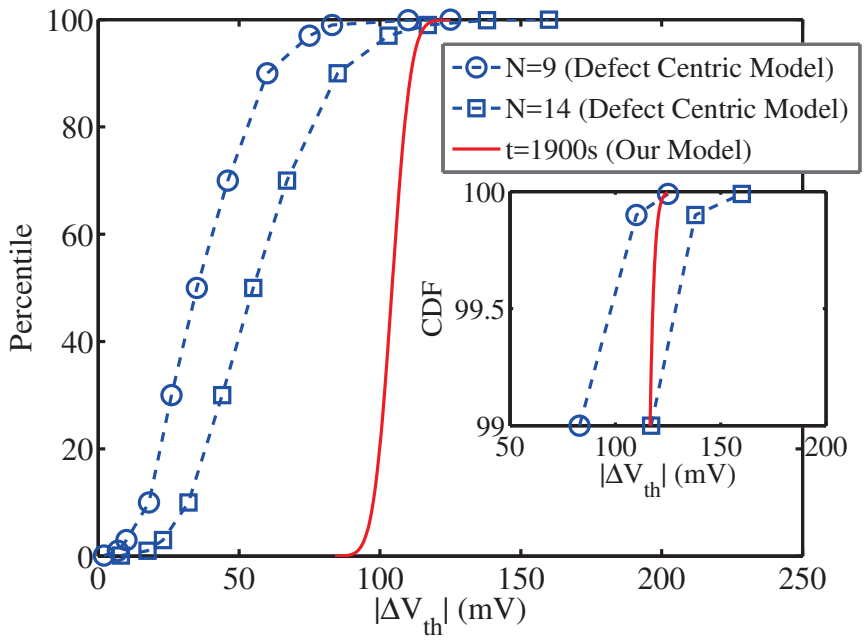
We have analyzed a five-stage ring oscillator in order to examine the effect of PBTI. We did HSPICE [45] Monte Carlo simulations on 1000 samples using Predictive Technology Nodes (PTM) [46]. Fig. 2.12 shows the plot of σ and change in μ with respect to the time-zero mean oscillation frequency for two different technology nodes. The circuits were subjected to a stress condition for 10^4 seconds with a constant stress of 2V and 125^oC. We observe an increase in variability and randomness with the decrease in device area. The inset shows the change in average oscillation frequency when the circuits are relaxed at 0V for 10^4 seconds. We see about 25% recoveries within that period of time which is similar to what we have observed in our MC simulation of individual devices in Fig. 2.7(b). In addition, the degradation at circuit level follows similar power law as that of the device level dependance.

2.4 Conclusion

In this chapter, we have proposed a statistical model to predict the effect of PBTI in nMOSFET lifetime and the variation was incorporated in HSPICE in order to see the PBTI impact on circuit delay. The simulation results show substantial random variation of performance in both device and circuit level. We have also observed that the variation becomes worse with the increase in stress condition (voltage and



(a)



(b)

Fig. 2.11. Comparison of our PBTI model to two different NBTI distributions: (a) Skellam function (2V, 140^{o}C) [43]; (b) Defect centric approach (2V, 125^{o}C) [44] for different number of defects $'N'$.

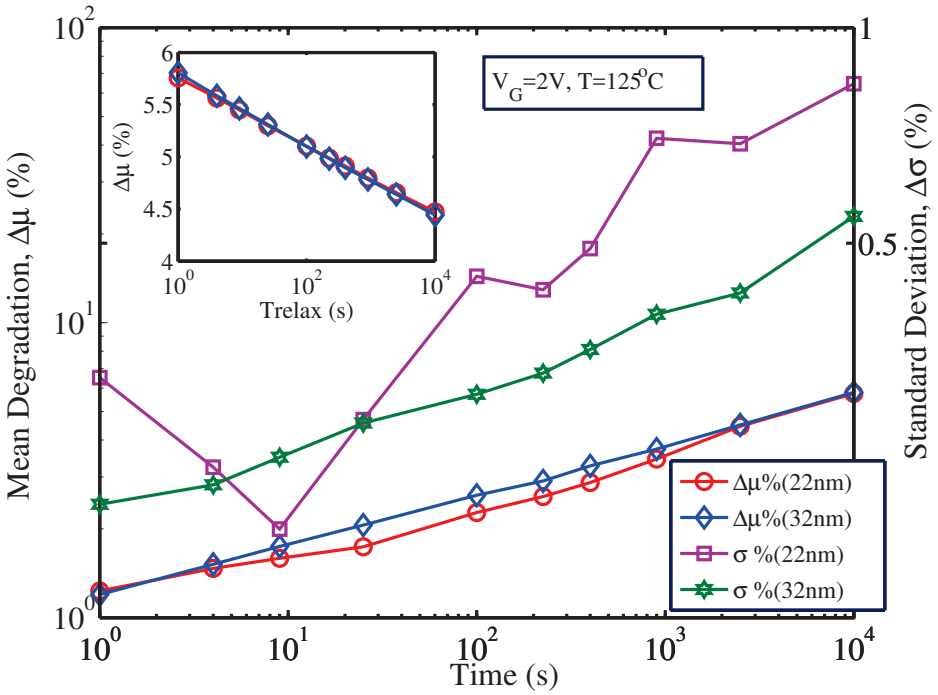


Fig. 2.12. Variation of μ_f and σ_f for 22 and 32nm PTM. The inset shows the recovery effect of the mean oscillation frequency.

temperature) and decrease in device size. Our model is scalable and can reproduce the experimental data provided by different research groups. With the statistical distribution predicted by this model, the impact of PBTI can be analyzed at the circuit level of design abstraction.

3. INVESTIGATION OF DEPENDENCE BETWEEN TIME-ZERO AND TIME-DEPENDENT VARIABILITY IN HIGH-κ NMOS TRANSISTORS

The time-independent process-induced variation has increased because of the aggressive scaling down of devices. As a result, the faster devices with lower threshold voltage distribution tail experience higher stress, leading to additional skewness in Bias Temperature Instability (BTI). Since the same traps in NMOS devices contribute to Positive BTI (PBTI), Time-Dependent Dielectric Breakdown (TDDB), and Stress-Induced Leakage Current (SILC), it is necessary to investigate the effect of time-zero variability on all of these effects simultaneously. Accordingly, we propose a simulation framework to model and analyze the impact of time-zero variability (in particular, random dopant fluctuations) on different aging effects. For small area devices ($\approx$1000 nm^2) in 30nm technology, we observe significant effects of Random Dopant Fluctuation (RDF) on BTI-induced variability ($\sigma\Delta V_{th}$). In addition, circuit analysis reveals a similar trend in performance degradation. However, both TDDB and SILC show weak dependence on RDF. We conclude that the effect of RDF on V_{th} degradation may not be disregarded in scaled technology and needs to be considered for variation-tolerant circuit design.

3.1 Introduction

Both time-dependent and process induced variations play critical roles in reliability assessment at both device and circuit levels [12, 47]. In addition, the dependence between time-zero and time-dependent variability needs to be properly addressed. The devices at the lower process-induced distribution tail performs faster (Fig. 3.1) due to their low turn-on voltage compared to the nominal threshold voltage, $V_{th}(0)$.

The variation can be up to few sigma values ($\sigma_(V_{th}(0))$) of the time-zero variation. Since the amount of carriers inside the channel region is influenced by the process variations (i.e. there are more carriers in the faster devices) [48], the trap generation rate inside the high-κ and interfacial layer (IL) is expected to be higher in those faster devices. As an example, in order to comprehend the nature of RDF and temporal BTI evolution and their dependence, it is important to capture the effect of each dopant in the channel region and defect in the bulk oxide layers simultaneously and determine the corresponding effect on V_{th} degradation. Moreover, PBTI, TDDB, and SILC are correlated in NMOS transistors [14]. Therefore, a complete modeling framework of RDF as well as BTI variation–considering the fluctuation in number and position of independent dopants and their impact on oxide defects/traps is necessary for circuit analysis. The rest of the Chapter is organized as follows. In section 3.2, we explained our modeling and simulation framework. In this section, we have extended our PBTI model [49] in order to analyze TDDB and SILC. Section 3.3 discuses the possible dependence between RDF and BTI at the device and circuit levels of design. We have also carried out TDDB and SILC analysis for NMOS devices considering this interrelation. Finally, section 3.4 concludes this work.

3.2 Framework for Variability Analysis

Fig. 3.2 shows the framework of our analysis. We have investigated the dependence between time-zero variation (RDF) and time-dependent variability and evaluated its impact on aging effects such as BTI, TDDB, and SILC. Understanding and modeling such an interrelation can be significant in reliability-aware circuit design. In order to properly model the impact of RDF on different aging effects, the non-uniformity of the electric field across the dual oxide layers (HK and IL) along the channel length needs to be considered. To that effect, we have used a TCAD simulator to carry out our device-level simulation [15]. We have extracted the electric field as a function of the channel length. The extracted fields are then incorporated to

our BTI trap generation model [49] and calibrated [14]. Subsequently, we integrated our PBTI model to the device simulator in order to analyze the dependence between RDF and PBTI. The TDDB and SILC analyses are carried using our extended trap generation model developed in MATLAB [50].

The NMOS device designed in the TCAD simulator is benchmarked with ITRS 2009 at the 30nm technology node [7]. The thicknesses of the high-κ and interfacial layers are 2nm and 0.6nm, respectively. Other design parameters are shown in Table 3.1. The ON and OFF currents of the simulated NMOS device are 1190 $\mu A/\mu m$ and $100nA/\mu m$, respectively. We have used the drift-diffusion transport model and the unified mobility model from Philips for the simulations [51]. For quantum correction, the device simulator uses the Van Dort model [52]. A detailed explanation of these models and a list of all the parameters used are available in the device user guide of the TCAD simulator [15]. Fig. 3.3 shows the cross-section of our NMOS device.

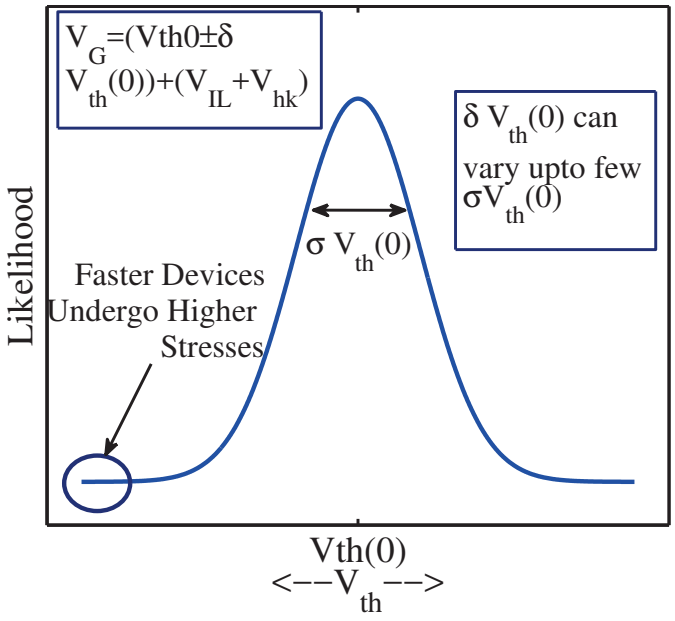


Fig. 3.1. Time-zero variation $\delta_{V_{th}(0)}$ with respect to the nominal threshold voltage, $V_{th}(0)$ can be up to a few $\sigma_{V_{th}(0)}$. This can cause a non-uniform stress voltage distribution across the oxide layers.

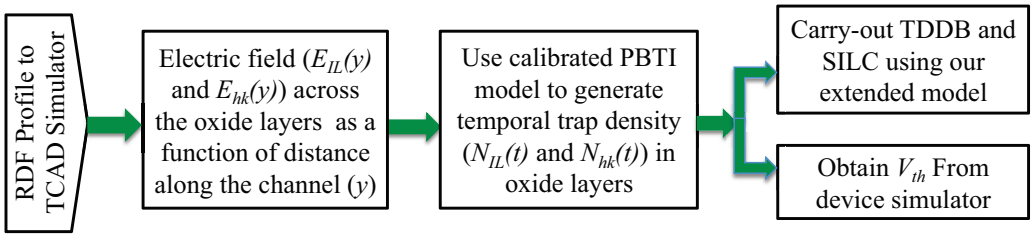


Fig. 3.2. Our Simulation Framework.

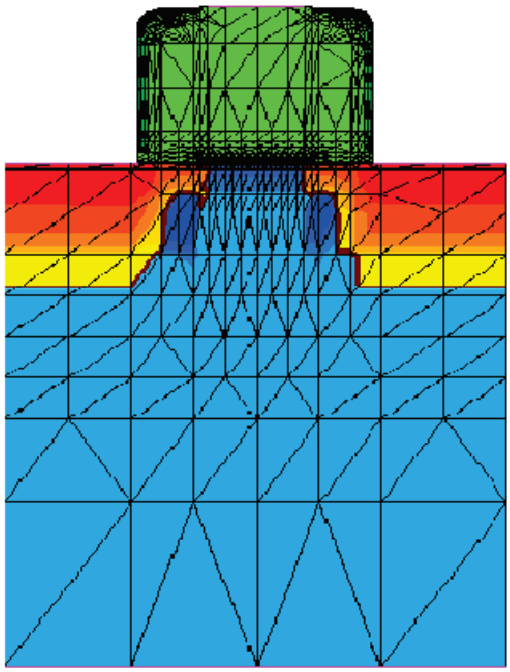


Fig. 3.3. Cross-section of our simulated NMOS device.

Unless otherwise specified, the stress voltage and temperature were maintained at 2V and 125^{o}C. We applied a drain voltage of 50mV in the device simulator and extracted V_{th} using the maximum-transconductance method, which is a built-in feature of the TCAD simulator. The various parameters used in our reliability analysis are shown in Table 3.2. Below we discuss RDF modeling and analysis and different time-dependent aging effects integrated into our simulation framework.

Table 3.1.
NMOS and PMOS Design Parameters

Parameter Name	**Value**
Gate Length (L)	30nm
Gate Width (W)	35nm
Gate Oxide Material	HfO_2, SiO_2
Equivalent Oxide Thickness	0.95nm
Channel Doping	$2.7x10^{18}/cm^3$
Source/Drain Doping	$6x10^{20}/cm^3$
Halo Doping	$2x10^{19}/cm^3$
Source/Drain Extension	10nm
Spacer Length	20nm
NMOS Gate Workfunction	4.6eV
PMOS Gate Workfunction	5.06eV

3.2.1 Time-zero Variability

Time-zero fluctuation is comprised of three major components (1) Random Dopant Fluctuation (RDF), (2) Mean Gate Length Deviation (GLD), and (3) Line Edge Roughness (LER) [1]. Among these three, RDF is considered to have the most adverse effects [1, 4, 53]. With the aggressive scaling down of MOSFETs, the intrinsic variation of $\sigma Vth(0)$ due to the smaller number of discrete implanted dopants and their corresponding random position in the channel becomes more distinct [4, 54]. RDF causes a large variation among similar transistors and affects performance at the device level as well as the circuit level of design abstraction [55]. In this work, we have generated RDF profiles considering randomness in the number of dopants using a pseudo-random number generator [50].

$$N^i_{RDF(0)} = Poiss(N_{RDF_{avg}}(0))_{1XS}; i = 1, 2, ..., S. \tag{3.1}$$

Table 3.2.
Parameters used in our simulation

Parameters	Value	Reference	Parameters	Value	Reference
$\alpha(HfO_2)$	0.17	[8, 14]	$\alpha(SiO_2)$	0.38	[56]
$E_r(HfO_2)$	0.8eV	[57]	$E_r(SiO_2)$	0.1eV	[58]
$a_0(HfO_2)$	1nm	[31]	$a_0(SiO_2)$	0.6nm	[31]
Φ_B	3.2eV	[59]	σ_n	$10^{-14}/cm^2$	[16]
a_0 : Trap size σ_n : Trap cross-section E_r : Trap relaxation energy α : Power exponent of time dependence Φ_B : Barrier height of the electrons					

where, $N^i_{RDF(0)}$ is a randomly drawn number from an 1xS vector generated by Poisson′s distribution. $N_{RDF_{avg}}(0)$ is the average doping density in the channel region and S is the number of microscopically similar sample devices. In addition to the number of random dopants, randomness of the dopants' location also plays a significant role in V_{th} variation. Dopants that are closer to the Si/SiO_2 interface have a larger influence on the V_{th} fluctuation [54]. Moreover, Drain-Induced Barrier Lowering (DIBL) increases the standard deviation of V_{th} in nanoscale transistors that is attributed to RDF. This is because the drain potential can nullify the effect of dopants that are located close to the drain contact; a large number of dopants close to the drain contact can cause high potential profile [55]. The RDF profiles are fed to the TCAD simulator and the locations of the dopants are randomized using the built-in feature of the TCAD simulator.

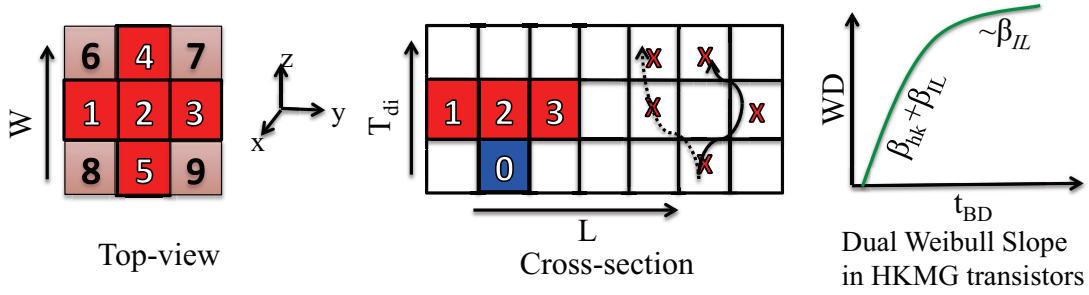


Fig. 3.4. Implementation of the 3D percolation model. One single defect in the bottom layer can lead to multiple breakdown paths. $'X'$ indicates a defect generated in a cell. Dual Weibull slopes confirm different defect generation rates in HK and IL layers.

3.2.2 Time-Dependent Dielectric Breakdown (TDDB)

The time-dependent dielectric oxide breakdown (TDDB) in HKMG transistors is characterized by short breakdown times as well as shallow Weibull slopes [31]. We have incorporated a 3D percolation model into our BTI model in order to capture the TDDB effect in the IL and the HK layers. Although the high-κ materials demonstrate lower tunneling currents and improved transistor performance, a thin interfacial layer (IL) of SiO_2 is needed to passivate the Si-dielectric interface [31]. Since the defect generation rate in SiO_2 at operating voltage is lesser than the high-κ layer [14], the presence of this IL layer improves both TDDB and PBTI and also increases the carrier mobility [14, 31].

The formation of conduction paths due to the accumulation of oxide traps triggers the breakdown of a dielectric layer. The widely-accepted percolation model [60] captures this idea for thick oxide layers. However, in scaled technologies, the formation of 3D conduction paths can no longer be ignored since the oxide layers are comprised of very few defect layers. Subsequently, a 3D cell-based percolation model has been proposed and implemented in different literature [31, 61]. In our statistical model, we have incorporated this model in-order to capture the TDDB effect. The Weibull plot is determined as:

$$W = ln(-ln(1 - F)) \tag{3.2}$$

where, F is the fraction of failed devices in a Monte-Carlo (MC) event. Fig. 3.4 shows the cross section and top-view of a three-layer dielectric material and the dual Weibull slope observed in HK gate stacks. We have considered all of the nearest neighbors in determining a possible link between the two adjacent layers. In Fig. 3.4, a defect $'0'$ in the bottom layer can have nine possible connections to the layer above itself. Defects $'1'$, $'2'$, $'3'$, $'4'$, and $'5'$ have at least one common edge with $'0'$, while defects $'6'$, $'7'$, $'8'$, and $'9'$ have exactly one node in common with $'0'$. When a continuous path is established between the top and bottom layers, a conduction or breakdown path is considered to have formed. In the case of the dual-layer stack that is widely used in today's technology, TDDB analysis has become more complicated

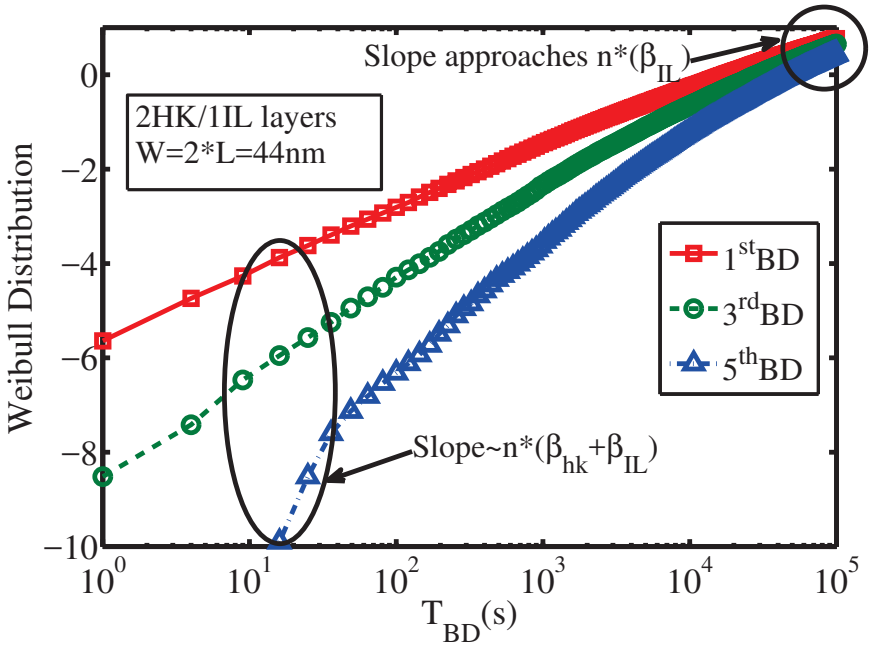


Fig. 3.5. Dual-Weibull slope is observed due to the inhomogeneous defect generation rates in the HK and IL layers [31].

since high-κ and interfacial layers have heterogeneous defect generation rates. High-κ layers are prone to the formation of both native and stress-induced defects. Hence, they wear out relatively faster than the IL layer, and the ultra-thin IL layer determines the final catastrophic breakdown [62]. In order to validate the model, we applied a very high stress condition and ran our MC simulation with 15000 samples for 10^5 seconds. We observe from Fig. 3.5 that the slopes for different breakdowns, although initially dependent on defect formation in both HK and IL layers, gradually become independent of the HK layer as the stress time increases. The Weibull slope β in the dielectric layers is defined as [31]:

$$\beta = \alpha(\frac{T_{ox}}{a_0}) \tag{3.3}$$

where, α is the time exponent for defect generation, a_0 is the cell or defect size, and T_{ox} is the thickness of the either HK or IL layer. We observe in Fig. 3.5 that the Weibull plot slowly approaches the slope in the IL layer since the final hard breakdown of the device is dictated by the trap generation in the IL layer. However, the actual reason behind the bi-modal distribution observed by different groups is still controversial. Yew et al. [63] demonstrated that the presence of Grain Boundary (GB) defects in the HK layer can cause early breakdown in many devices and can give rise to a similar bi-modal Weibull distribution. We have also considered this effect in our simulation results and discussion section.

3.2.3 Stress-Induced Leakage Current (SILC)

SILC is a means to analyze the buildup of oxide traps under BTI stress. As the trap density reaches a critical value, the SILC current can cause significant power consumption, leading to the hard breakdown of a device [64]. Therefore, it is necessary to understand the evolution of SILC and have appropriate model that can predict the progressive wear of the oxide layers in a scaled technology. Fig. 3.6 shows the band diagram of an NMOS under inversion mode and the components of gate leakage currents. There are three major components of gate leakage [27]:

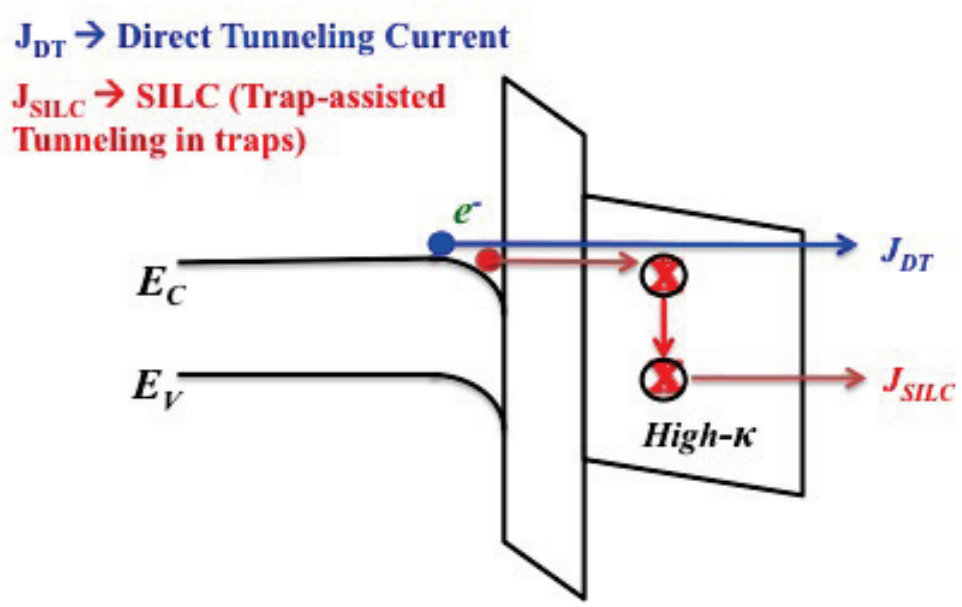


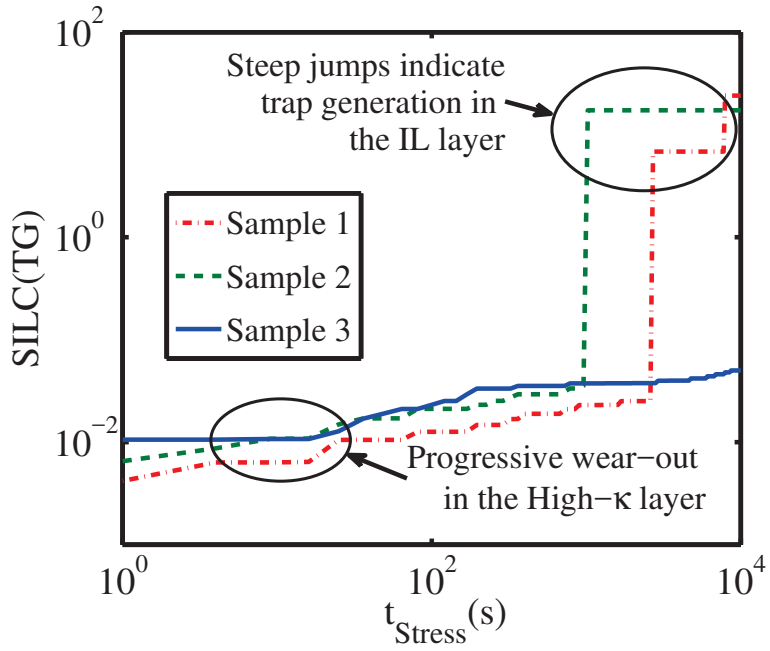
Fig. 3.6. Trap-assisted SILC in HK layer decreases due to high relaxation energy.

- Direct Tunneling Current (J_{DT}) or the time-zero gate leakage
- Trap-Filling Current in preexisting traps (J_{TP})
- Stress-Induced Leakage Current (J_{TG}) through Trap-Assisted-Tunneling in generated traps

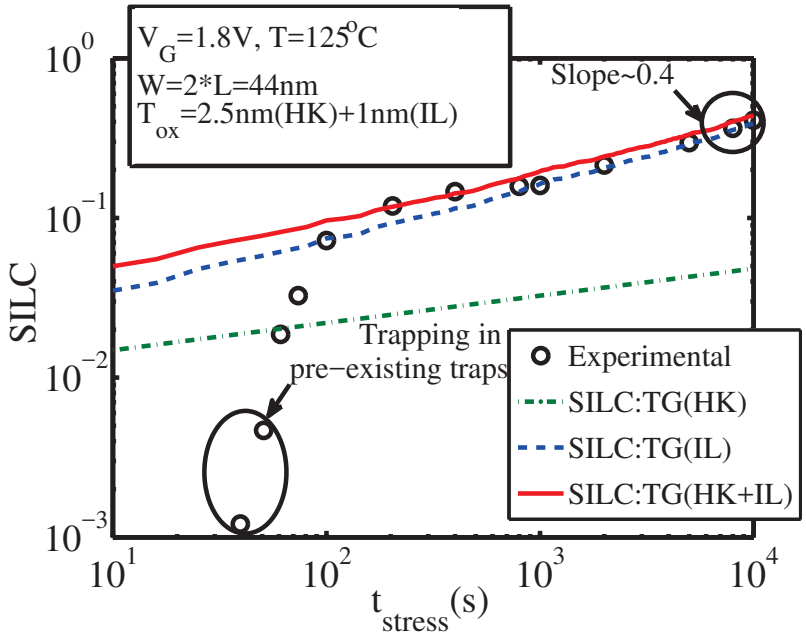
The time-zero leakage current is known as direct tunneling current J_{DT}, while the stress-induced leakage current J_{SILC} is known as Trap-Assisted-Tunneling (TAT) current. J_{SILC} increases when a device is under electrical stress. We implemented the modified direct tunneling current for a high-κ gate stack proposed in [59]:

$$J_{DT} = \frac{q^3}{8\pi h\phi_B}\frac{\phi_B}{V_{SiO_2}}\left(\frac{2\Phi_B}{V_{SiO_2}} - 1\right)\frac{V_{SiO_2}}{t_{SiO_2}}T_{WKB} \tag{3.4}$$

where, Φ_B is the barrier height seen by the carriers in the conduction band of the substrate, V_{SiO_2}, t_{SiO_2}, and T_{WKB} are the voltage drop across the SiO_2 layer, thickness of the IL layer, and the tunneling probability across the dual gate stack determined



(a)



(b)

Fig. 3.7. Comparison of: (a) SILC in different small area sample devices indicates that trap generation in IL layer dominates total SILC; (b) Average SILC in 15000 such devices matches the experimental data [14].

using the WKB approximation [16], respectively. The stress-induced current due to trap generation (TG) is modeled as [14]:

$$J_{TG} = \int\int k\sigma_n V_{th} N(t) T_r (f_c - f_a) dx dE \tag{3.5}$$

Where,

$$T_r = \frac{T_{in}(E) T_(out)(E - E_r)}{T_{in}(E) + T_{out}(E)}$$

where, T_{in} and T_{out} are the tunneling probabilities to and from the traps, respectively; E_r is the trap relaxation energy, f_c and f_a are the occupation factors of cathode and anode, respectively; σ_n is the electron capture cross section of the traps, and v_{th} is the thermal velocity. k in (3.5) depends on the electron concentration of the conduction band and the energy suppression factor. In our work, we have calibrated k in order to fit the experimental data in [14]. The normalized SILC current is then defined as [65]:

$$J_{SILC}(t) = \frac{J_{TG}(t) - J_{TP}(t_{sat})}{J_{DT}(0)} \tag{3.6}$$

where, t_{sat} is the saturation time of pre-existing trap filling, which is usually in sub-milliseconds [11], and J_0 is the gate leakage current at t=0, which includes the direct-tunneling component and tarp-assisted tunneling in the pre-existing defects. The actual reason behind SILC is a controversial issue. While some groups attribute SILC phenomenon to the defect generation in the HK layers [27,66,67], others explain SILC in terms of defect generation in the IL layer only [62, 68]. J. Yang et al. [14] argued that SILC is caused by defects in both HK and IL layers, but mostly dominated by IL layer defects with much lower relaxation energy. In our model we have considered degradation in both HK and IL layers and calibrated our model accordingly [14]. Fig. 3.7 shows the average SILC current for a MC simulation of 15000 similar devices. As we can observe from Fig. 3.7(a), for different small area sample devices, total SILC due to trap generation spikes when there is defect generation in the IL layer. This is because the relaxation energy in an IL defect is much lower than that of an HK defect [14]. The higher relaxation energy in the HK layer defects can exponentially decrease

the tunneling out (T_{out}) probability and significantly decrease the contribution to the overall SILC. As a result, the slope of the average SILC (Fig. 3.7(b)) gradually approaches the time exponent of trap generation in the IL layer. We have ignored the contribution from pre-existing traps and assumed J_{TP}=0 for our analysis since the trap density is not provided in [14]. Besides, these technology-dependent native traps can be optimized to a very low density [26], in which case only the trap generation plays a role in the SILC measurements.

3.3 Results and Discussions

In this section, we explain the impact of RDF on time-dependent degradation using the results of our simulation framework. We extracted the electric field from the TCAD simulator. Fig. 3.8 shows the electric field variation across the IL layer among different devices with an average doping density of 2.7x10^{18}/cm^3 in the channel region at V_G=1V. We have divided the channel region in 35x30 number points and extracted the oxide electric field as a function of the channel length at each of the points along the width. We see variations in the oxide electric field due to the randomization of dopant locations as well as the numbers. Since trap generation is a function of the oxide electric field, we expect to see the impact of RDF in stress induced trap density profiles among a large number of microscopically similar devices. We incorporated the extracted electric field in our PBTI model [49] in order to further investigate the impact of RDF on different aging effects. In Fig. 3.9, we have plotted the local gate injection current as a function of channel length and width. The local injection current is determined using the direct tunneling current model in equation (3.4), and we plotted the normalized value (normalized with respect to the maximum value) for gate voltages of 1 and 2V in Fig. 3.9(a) and 3.9(b), respectively. For this simulation, we have considered the dopant profile corresponding to Fig. 3.8(a), where 36 dopants are randomly assigned inside the channel. We can see that RDF causes variation in the local injected current at both 1V and 2V. In order to investigate the relative

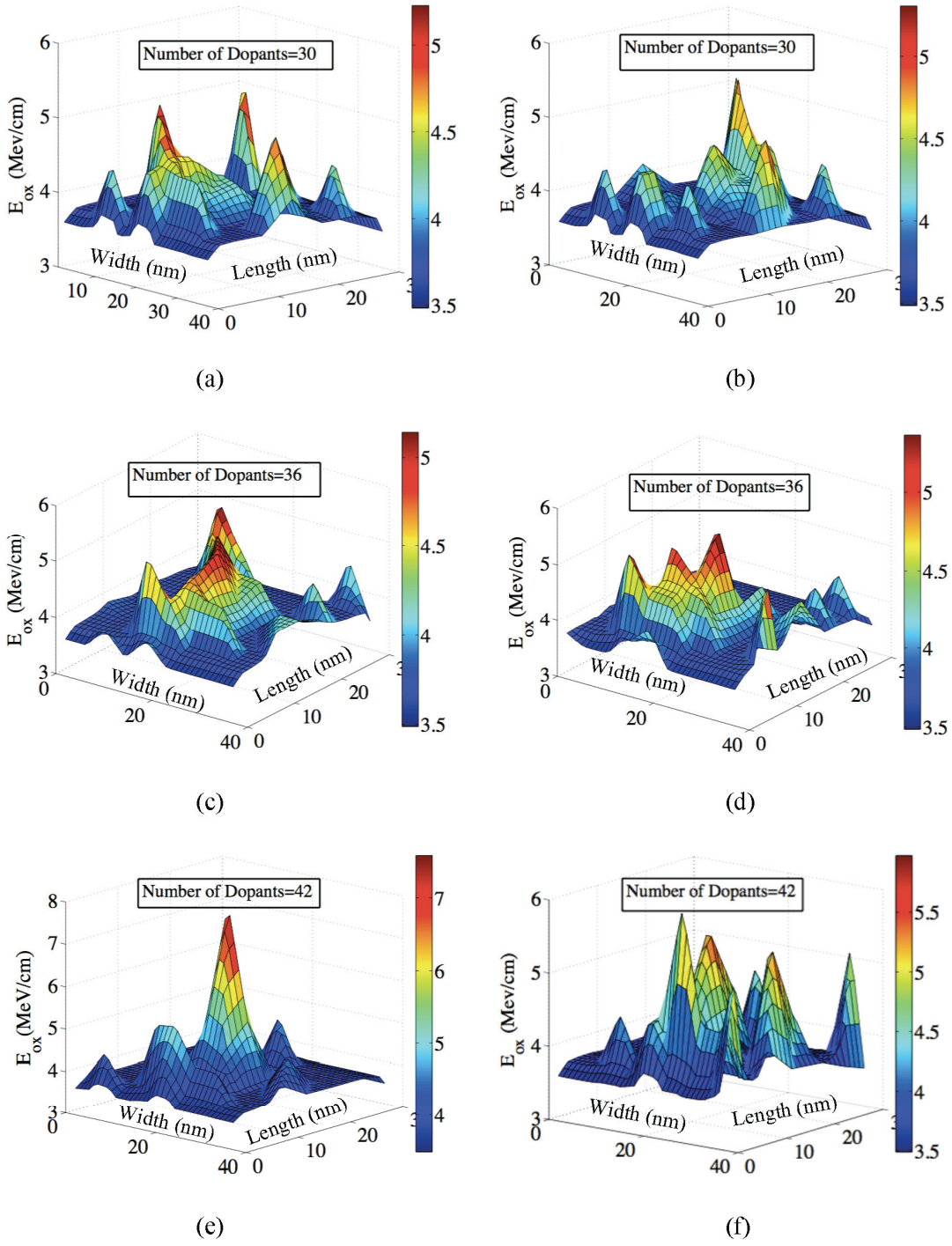
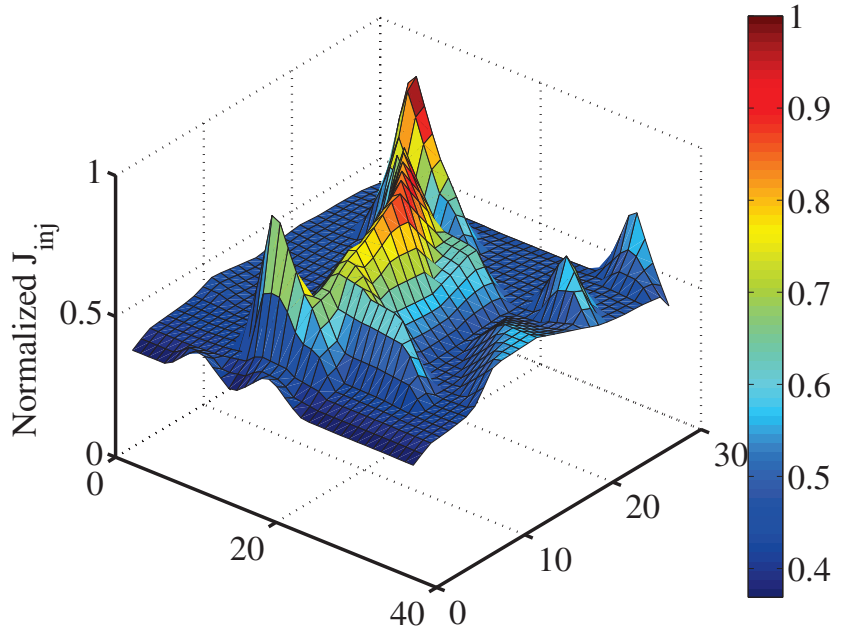
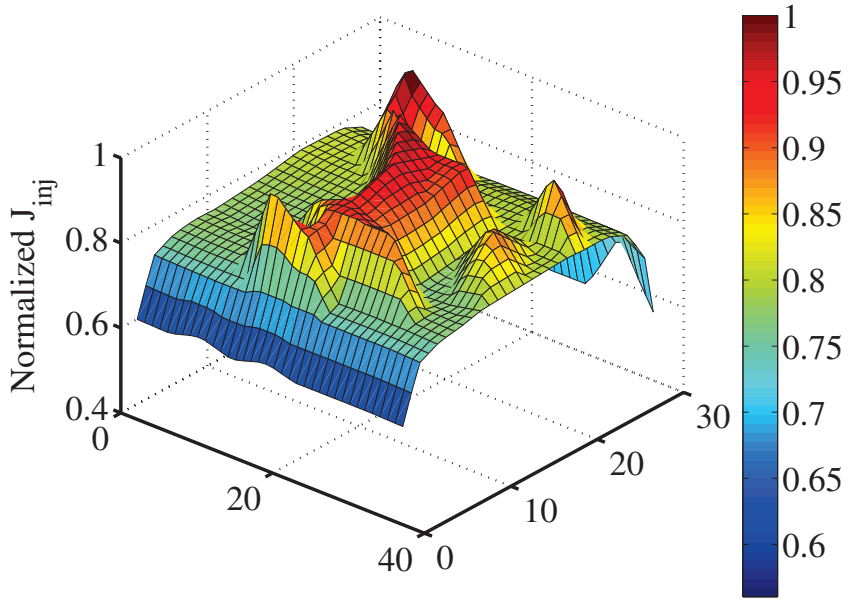


Fig. 3.8. Effect of random dopant fluctuation on the electric field across the IL oxide layer: (a), (b), and (c) show the electric field variation due to the fluctuation of the number of dopants in the channel region of an NMOS device. In comparison to the upper three figures, (d), (e), and (f) show the variation due to the fluctuation in dopant locations. A gate voltage of 1V is applied in these simulations.



(a)



(b)

Fig. 3.9. Normalized distribution of local injection current with 36 dopants inside the channel region: (a)V_G=1V; (b) V_G=2V. The effect of RDF is higher at a lower gate voltage.

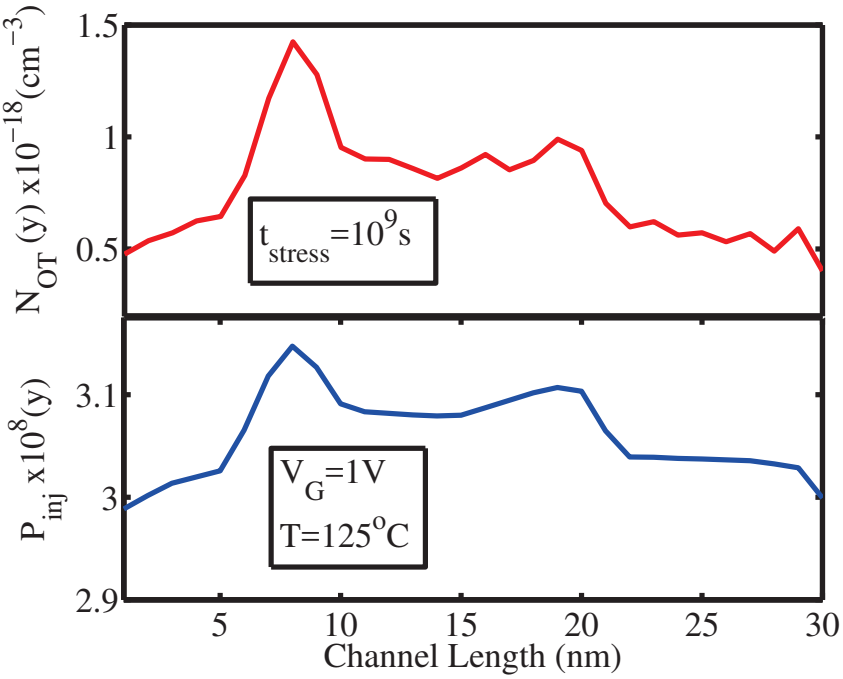


Fig. 3.10. Correlation between local injection probability (P_{inj}) and PBTI trap density (N_{OT}) along the channel at V_G=1V. The results indicate direct correlation between RDF and PBTI.

impact of RDF, we determined the standard deviation of the local injected currents as a percentage of their average values. We observe that the variability induced by RDF at 1V (21.9%) is much higher compared to that at 2V (10.3%). Therefore, the impact is expected to be more significant at the device operating voltage. Since PBTI is driven by the gate injection current, we investigated its correlation to local injected current probability (P_{inj}) in Fig. 3.10. For V_G=1V (corresponding to Fig. 3.9(a)), we extracted the probability as a function of the channel length at a random location along the width and incorporated the corresponding electric field distribution into our PBTI model. We simulated a large area device (10000x30 nm^2) for a stress period of 10^9 s and plotted the lateral distribution of the generated trap density in the upper panel of Fig 3.10. In the lower panel, we have plotted the corresponding local injection probability and observe that they are closely correlated. Therefore,

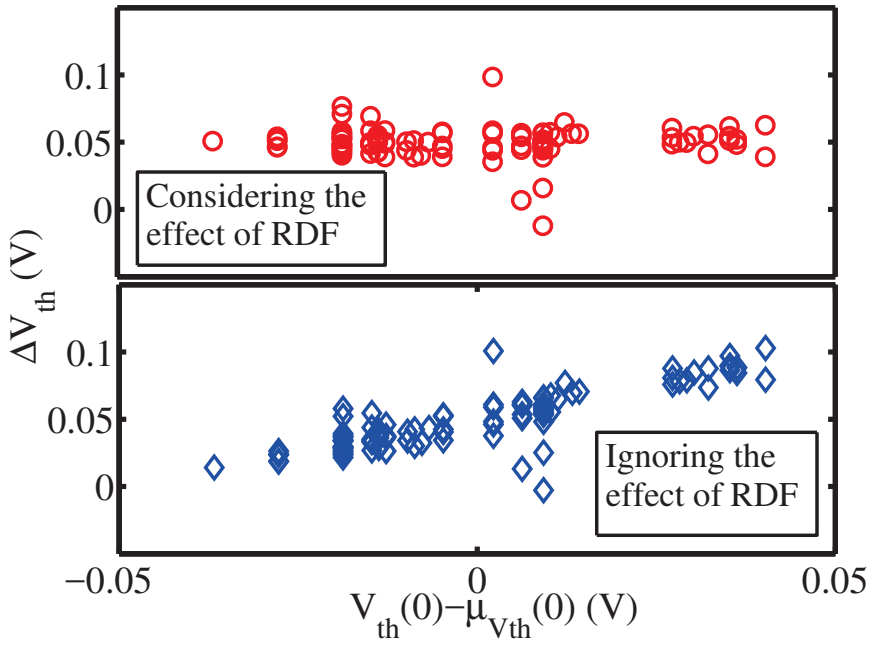


Fig. 3.11. Effect of RDF on PBTI at a mean ΔV_{th} of 50mV. Ignoring the effect of time-zero variability leads to erroneous results in BTI estimation. The simulation was carried out at 2V and 125^oC.

we conclude from Fig. 3.9 and 3.10 that the impact of RDF can modulate the δV_{th} distribution induced by PBTI.

Fig. 3.11 demonstrates the impact of RDF on PBTI estimation for 50mV of mean δV_{th} degradation. The lower panel represents the data where the BTI degradation is estimated using an average time-zero threshold voltage; the upper panel shows the same data taking the effects of RDF into account. δV_{th} after a stress period of t_{stress} for these two panels are defined as:

$$\Delta V_{th}(t_{stress}) = \begin{cases} V_{th}(t_{stress}) - V_{th}(0), & \text{upper pannel} \\ V_{th}(t_{stress}) - \mu V_{th}(0), & \text{lower pannel} \end{cases}$$

Both sets of data are plotted against the difference between the time-zero threshold voltage, $V_{th}(0)$ for each of the devices and the mean of the time-zero threshold voltage, $\mu V_{th(0)}$. In this plot, the devices on the left side are relatively faster, and we see that

δV_{th} for these devices are higher in the upper panel. The effect of RDF on the devices that have threshold voltages close to the mean $\mu V_{th}(0)$ are almost negligible. The devices on the right side of the panels, on the other hand, have higher time-zero threshold voltages. We see that the δV_{th} in the upper panel is relatively less than that of the lower panel. In other words, RDF causes higher degradation to faster devices and lesser degradation to slower devices, as we have explained in Fig. 3.1.

Fig. 3.12 shows the impact of RDF on the threshold voltage degradation at 25mV, 50mV, and 75mV of PBTI degradation. $\sigma_{PBTI}(V_{th}(0))$ is the PBTI standard deviation, taking the effect of RDF into account, while $\sigma_{PBTI}(\mu V_{th}(0))$ represents PBTI variability considering average time-zero V_{th}. We defined a dependence parameter ζ_D, which is the underestimation of $\sigma_{\Delta Vth(PBTI)}$ with respect to mean $\Delta Vth(PBTI)$ when the effect of RDF is ignored and plotted on the right y-axis of Fig. 3.12. We observe that the RDF effect on the threshold voltage degradation decreases with increased BTI effects. This trend follows the experimental data presented in [12]. This observation can be significant for scaled technology where the margin for parametric variation is very stringent [69].

Next, we investigated the effect of RDF on TDDB and SILC at a stress voltage of 2V. The analysis was carried out for 15000 sample devices. Fig. 3.13 shows the Weibit plot for both the 1^{st} and the 3^{rd} breakdown with and without considering the effect of RDF. We observe negligible differences in the TDDB lifetime. We carried out similar analysis, taking the presence of GB defects into consideration. We divided the HK oxide layer in 42 grain-regions and assigned an average pre-existing GB defect density of $10^{19}/\text{cm}^3$ in 25% of the sample devices. Since the presence of GB defects compromises the underlying IL layer [63], we assumed a 10 times higher IL layer defect generation rate in these devices. The simulation results for the 3^{rd} soft breakdown is plotted in Fig. 3.14, and we observe the negligible effect of RDF on TDDB. Early breakdowns due to GB defects in some of the devices give rise to a bi-modal Weibull distribution, which is consistent with the observations reported in [63]. Since TDDB

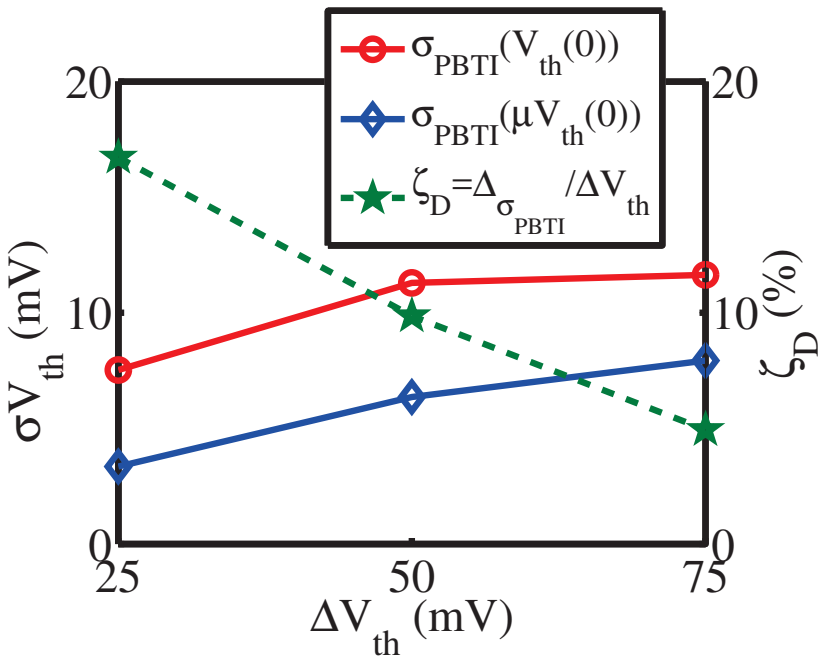


Fig. 3.12. Effect of RDF on PBTI decreases with increased BTI degradation.

requires higher stress conditions and defect density than BTI, the effect of RDF is negligible in both Fig. 3.13 and 3.14.

A similar trend is observed for SILC analysis and shown in Fig. 3.14. We observe almost indistinguishable SILC for both average doping effect and RDF. SILC is dominated by the trap generation in the IL layer, where the defect generation rate is assumed to be very small in our simulations. Due to the low density of generated traps in the IL layer, the statistical fluctuation due to the RDF effect has a trivial effect on SILC. SILC is highly a controversial issue, and the relative contribution of stress-induced traps in the HK and IL layers is still debatable. Therefore, further investigation on the effect of RDF on SILC is necessary.

Finally, we investigated the effect of RDF on BTI degradation at the circuit level and analyzed the impact of RDF-BTI correlation on the degradation in the frequency of oscillation (f_o) and transient supply current (I_{DDT}) of ring oscillators. For this analysis, we have considered both PBTI in NMOS transistors and Negative Bias

Table 3.3.
NBTI model and the parameters used in our simulations

The average number of interface traps generated is predicted by the following $H - H_2R - D$ model [70]:

$$N_{IT_{avg}}(t) = WL(K_H/K_{H_2})^{1/3}(k_f(y)N_0/k_r)^{2/3}(6DH_2t)^n \qquad (3.A)$$

$$k_f(y) = k_{f0}p_h(y)\exp(\gamma E_{ox}(y))\exp(-\frac{E_{AF}}{qV_T}) \qquad (3.B)$$

where, n, γ, k_f, k_r, N_0, k_H, k_{H_2}, and DH_2 are the NBTI power exponent, field acceleration factor, Si-H bond breaking rate, Si-H bond annealing rate, the pre-stress Si-H bond density at the Si/SiO_2 interface, generation rate, dissociation rate, and diffusion coefficient for H and H_2, respectively. y, W, and L in (3A) are the distance along the channel length, width, and length of transistors, respectively. p_h and E_{ox} are the hole concentration inside the channel region and the total oxide electric field, respectively. k_{f0} is the calibration parameter of our model.

If we assume $X=\{k_r, k_H, k_{H_2}, D_{H_2}\}$, X can be expressed by the following equation:

$$X = X_0\exp(-\frac{E_{A_X}}{qV_T}) \qquad (3.C)$$

E_{A_X} are the activation energies and $X_0=\{k_{r0}, k_{H0}, k_{H_20}, D_{H_20}\}$ are constants.

Parameters extracted from [70]:

$n=\frac{1}{6}$, $\gamma = 0.6 \pm 0.05$, E_{AF}=0.36eV

Parameters extracted from [71]:

$E_{A_{k_r}}$=0.2eV, $E_{A_{k_H}}=E_{A_{k_{H_2}}}$=0.3eV, $E_{A_{k_f}}$=0.175eV, $E_{A_{D_{H_2}}}$=0.58eV, k_{r0}=9.9x10^{-7}, D_{H0}=9.56x10^{-11} cm^2/s, k_{H0}=8.56 cm^3/s, k_{H_20}=5.7x10^5/s, D_{H_20}=3.5x10^{-5} cm^2/s

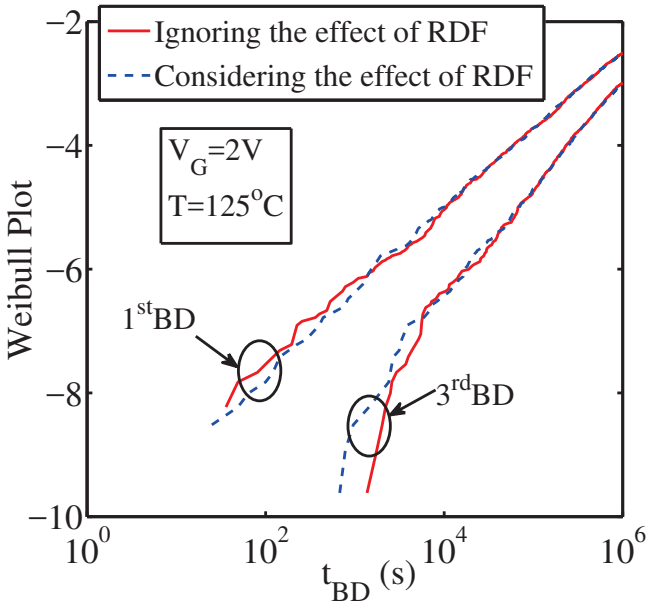


Fig. 3.13. The effect of RDF has very negligible or no impact on TDDB estimation.

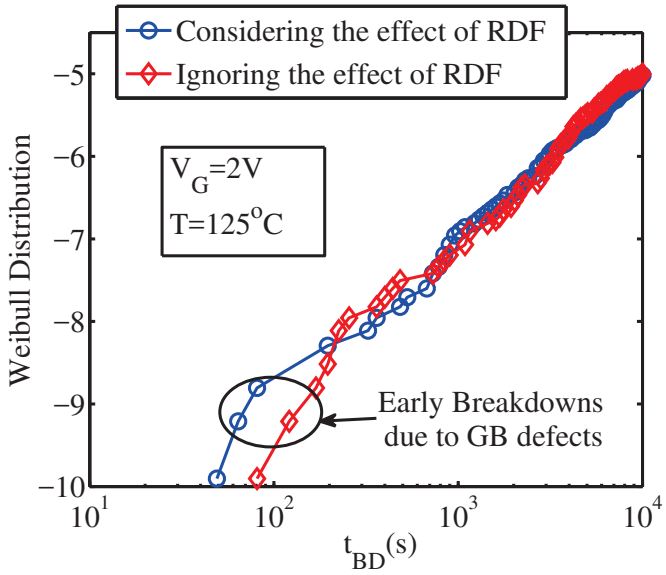


Fig. 3.14. Impact of RDF on TDDB, taking GB defects into account. The effect of RDF on TDDB is still negligible.

Temperature Instability (NBTI) in PMOS transistors. H-passivation is widely used to satisfy the dangling bonds at Si-SiO_2 interface and NBTI is manifested as V_{th} degradation due the hole-assisted dissociation of Si-H bonds at the interface (Fig. 3.16) [72]. We implemented the reaction-diffusion (RD) model of NBTI degradation [70, 72] and assumed Poisson′s distribution on the interface trap numbers among different microscopic samples [43]. The average number of interface traps was calibrated with the simulation results presented in [73]. The RD model used in our work is summarized in Table 3.3. We have used the same framework, as shown in Fig. 3.2, to take the effect of RDF on NBTI into account. We extracted $I - V$ characteristics from the TCAD simulator and implemented a look-up table based Verilog-A model for both n and p MOSFETs. Using the models, we carried out HSPICE [45] Monte Carlo simulations on 10000 samples. In order to explore the performance degradation

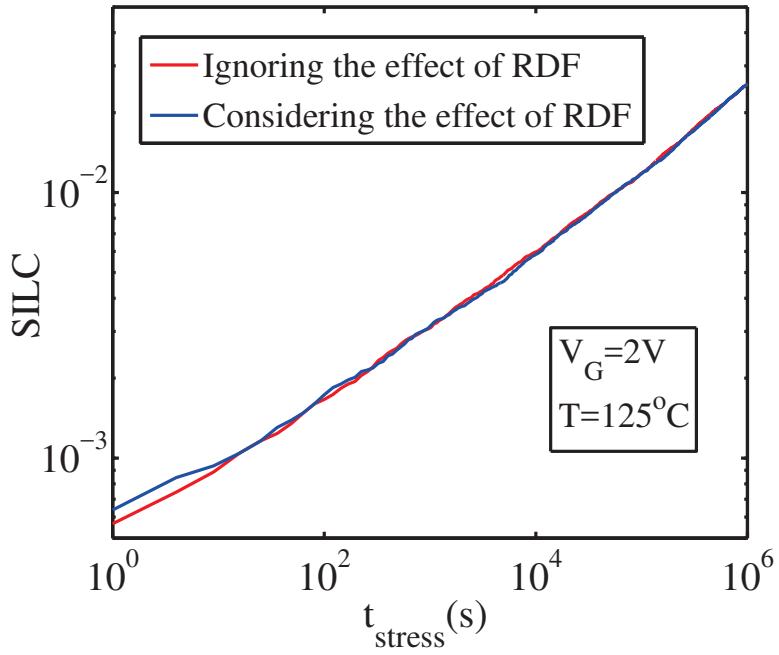
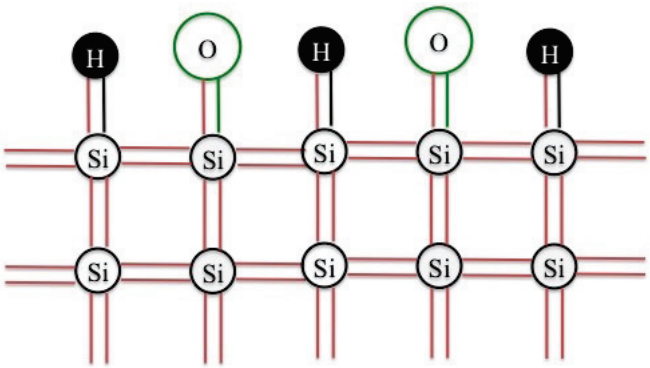
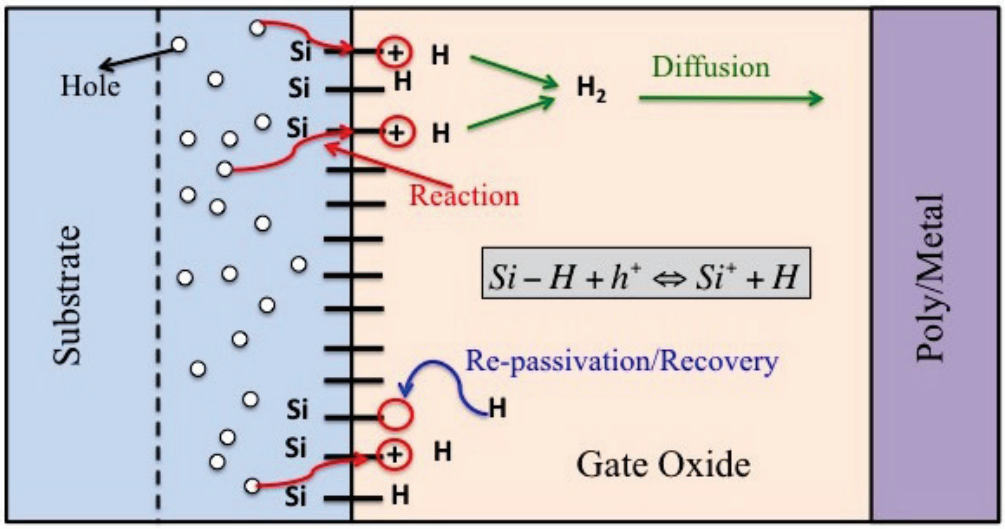


Fig. 3.15. Average SILC shows similar trend in TDDB since both are dominated by the trap generation rate of the IL layer.



(a)



(b)

Fig. 3.16. (a) H-passivation satisfies the dangling bonds at the Si-SiO_2 interface; (b) Hole-assisted breakage of Si-H bonds under NBTI stress [72].

of the relatively slower circuits, we determined the -3σ points of the above-mentioned parameters with and without considering the effect of RDF on BTI and plotted their difference as a percentage of the standard deviation (σ) in Fig. 3.17. We observe that ignoring the effect of RDF causes significant underestimation of both f_o and I_{DDT}, especially at a relatively lower BTI degradation. Similar to the device level results, the influence of RDF decreases as BTI degradation increases. In addition, we have

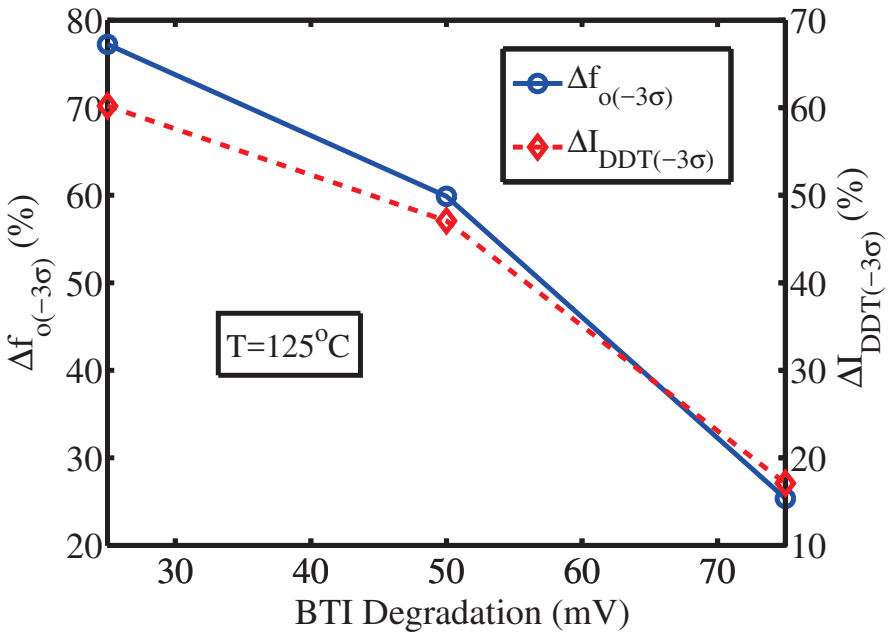


Fig. 3.17. -3σ points of frequency of oscillation and dynamic supply current as a percentage of their corresponding standard deviations. Effect of RDF on BTI effect is more noticeable at lower BTI degradation.

plotted the V_{DD} guard band required to recover the speed of the ring oscillators at the -3σ points in Fig. 3.18 for 5 and 7 stages ring oscillators. For this simulation, we have applied V_{DD} guard band so that the circuits at -3σ points can run at the same speed of the circuits with only RDF effect. We observe that the dependance of BTI on RDF decreases with an increase in the number of stages. The results suggest that the effect of correlation decreases with increased circuit complexity. Therefore, the effect of RDF on BTI may have to be considered for less complex circuits and scaled technology where the variability margin is strict.

3.4 Conclusion

In this Chapter, we have proposed a statistical modeling framework that takes the dependence between time-zero variability and various time-dependent aging effects

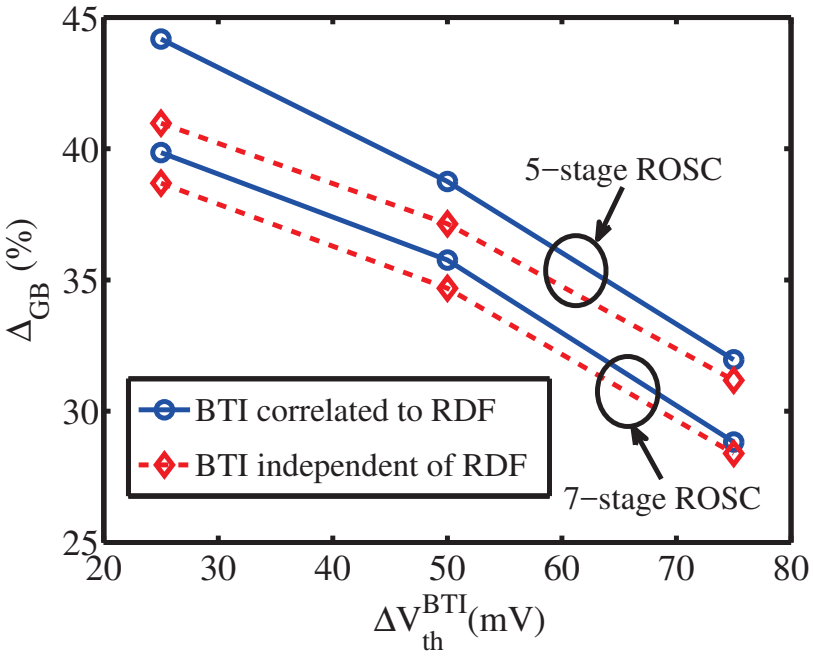


Fig. 3.18. V_{DD} guard band required to recover speed of ROSCs with 5 and 7 stages. The correlation is less noticeable in the 7-stage ROSCs.

into account. The framework is developed based on our proposed model of PBTI, a statistical RD model of NBTI, a 3D cell based percolation model of TDDB, and a TAT current model of SILC. We have analyzed small area devices (1000nm^2) and shown that: 1) RDF has considerable impact on BTI effect; and 2) the influence of RDF on time-dependent oxide breakdown is almost negligible. In addition, we carried out circuit analysis using a look-up table based Verilog-A Spice model and analyzed the effect of BTI degradation on the performance of ring oscillators. On both the device and circuit levels, we observed that the effect of RDF on BTI degradation is more pronounced at relatively lower BTI degradation. Since there are not enough experimental evidence that proves any impact of RDF on TDDB and SILC, further investigation is necessary. We conclude that the effect of RDF on V_{th} degradation needs to be properly addressed in scaled CMOS technology, and this may become crucial in the accurate guard band estimation of a circuit.

4. IMPACT OF THE COMBINED EFFECT OF RDF, NBTI, AND PBTI ON SRAM YIELD ANALYSIS

SRAM failure analysis is a significant component of microprocessor reliability analysis. In order to accurately estimate the stability of a SRAM array, it is necessary to take all variability and reliability effects into consideration. In addition, the dependence of aging effects such as BTI on process-induced RDF variability can make an accurate yield estimation even more challenging. We have considered both RDF and BTI effects simultaneously in SRAM failure analysis. We observe that the impact of BTI on RDF has a very negligible influence on worst case write operation. However, this dependence can have a considerable impact on SRAM worst case read and hold failure analyses, especially if the desired confidence level is 3-sigma (99.7%) or above.

4.1 Introduction

Static Random Access Memory (SRAM) occupies a large area of silicon semiconductor chips. In addition, their impact on the system performance plays a critical role in contemporary microprocessor reliability [74]. In scaled CMOS technology, process-induced variability as well as the time-dependent aging effects are considered two essential limiting factors of SRAMs [75]. In addition, SRAMs are driven to a 6 sigma confidence level [76]. Therefore, an accurate lifetime estimation of SRAMs is crucial in advanced technology. Besides, the scaling down of supply voltage in advanced technology has also compromised the robustness as well as reduced the read and write noise margins of memory cells. Reliability analysis of SRAM performance is one of the most researched topics in CMOS. The impact of NBTI, PBTI, and TDDB degradation on performance has been reported in literature to a great extent [40, 76–79]. However, none of these works take the impact of process-induced

variation on aging effects into account. In a conventional approach, the impact of RDF and BTI are considered uncorrelated and the combined effect is expressed by the following expression [12]:

$$\sigma_{Total}(t) = \sqrt{\sigma^2_{RDF}(0) + \sigma^2_{BTI}(t)} \tag{4.1}$$

where, σ_{RDF} is the time-zero variability due to RDF and σ_{BTI} is the time dependent variability due to BTI.

In Chapter 3, we have observed that Bias Temperature Instability (BTI) degradation is influenced by process-induced Random Dopant Fluctuation (RDF). Moreover, the impact of RDF on BTI is much higher at a nominal condition [80]. Therefore, it is necessary to estimate the combined effect of RDF and BTI at a relatively low stress condition (preferably at a nominal condition). For this work, we have leveraged the simulation framework proposed in Chapter 3. In section 4.2, we have briefly explained different failure mechanisms in 6T-SRAM cells. In section 4.3, we have carried out SRAM worst-case read, write, and hold time failure analysis with and without considering the effect of RDF. For our analysis, we have considered SRAM transient parameters, such as read and write access times, since dynamic stability analysis is preferable and gives a more accurate estimation of failure probability than the conventional Static Noise Margin (SNM) analysis [78, 81]. Finally, section 4.4 concludes the chapter.

4.2 Parametric Failure Mechanisms in an SRAM cell

Fig. 4.1 shows a 6T SRAM cell. There are four major parametric failure mechanisms that quantify SRAM cell stability [82]:

- Read Disturb Failure (Read Stability)
- Read Access Failure (Readability)
- Hold Failure (Data Retention)

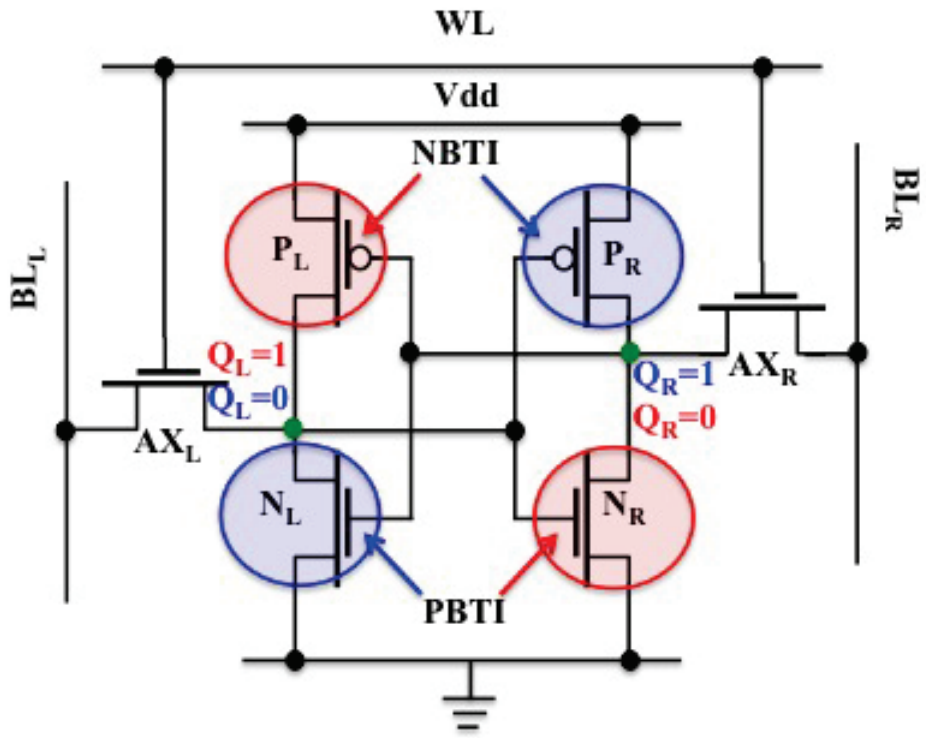


Fig. 4.1. A 6T SRAM Cell. Depending on value stored in the internal nodes, transistors $N_L - P_R$ (if Q_L=0, Q_R=1) or $N_R - P_L$ (if Q_L=1, Q_R=0) can be under BTI stress and become weaker over time.

- Write Failure (Writability)

We have briefly explained them in this section.

4.2.1 Read Disturb Failure

Read disturb occurs when the internal node storing logic $'1'$ discharges through various leakage paths and accidentally performs a write operation. During the read operation, both bit lines in Fig 4.1 are charged to V_{dd}, and depending on the stored value one of the bit lines discharges to $'0'$. During this process, the node storing $'0'$ can increase and accidentally flip the states of the nodes and thereby lose the stored data. As shown in Fig 4.2, if the word line pulse width is long enough to allow the node storing $'0'$ (e.g., Q_L in Fig. 4.1) to increase beyond the trip point

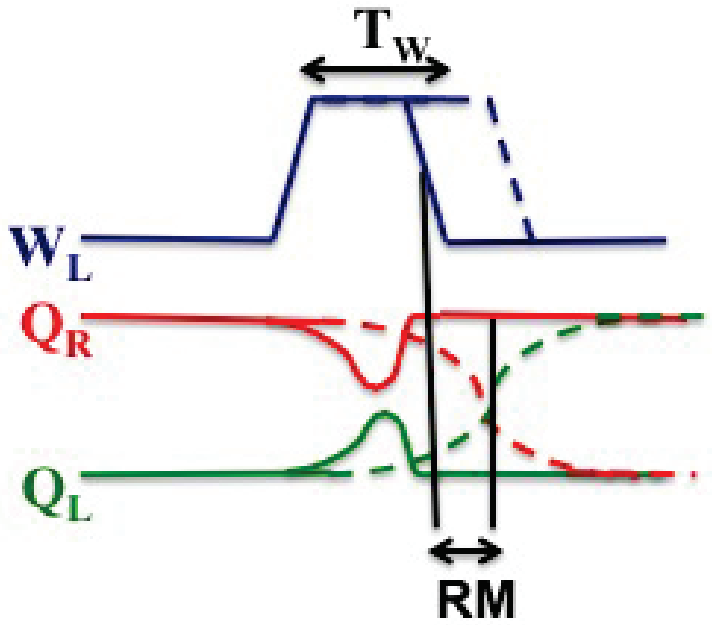


Fig. 4.2. Read disturb occurs if T_W is long enough to allow Q_L to discharge and flip the node states.

of the inverter $P_R - N_R$, the cell flips and causes a read failure. The dynamic read margin (RM) is defined as the time difference between the required time to cause a read failure and 50% of the time of the falling edge of T_W. As shown in Fig.4.2, this condition takes place if the word line pulse is wider than the required minimum time to perform the read operation. We have considered the worst-case read $'0'$ operation in our simulation.

Worst case read $'0'$ condition:

Node Q_L stores $'0'$ for t=t_1 where, t_1 is long enough to cause at least 10 mV mean BTI degradation to transistors P_R and N_L.

At t=t_1, read $'0'$ operation is performed.

Here, the assumption is transistors P_L and N_R are unaffected by BTI degradation.

4.2.2 Read Access Failure

In the case of the read operation, the bit lines are expected to produce a minimum voltage difference within a maximum tolerable limit (T_{Amax}). The time required to produce the voltage difference is known as the access time (T_{Access}). If the pull down and access transistors become weaker due to aging effects, cell access can take longer than T_{Amax}, and the condition is known as read access failure [82]. Due to PBTI, V_{th} of the access and pull down transistors can increase over time and cause read access failure. For example, if we consider that node Q_L in Fig. 4.1 is storing logic $'0'$, then transistors P_R and N_L are under BTI stress. If the value is stored for a long time, both transistors can become weaker. If T_{Access} becomes larger than T_{Amax}, the read operation fails to produce the desired voltage levels at the bit lines, therefore causing a read failure. In our analysis, we have considered the worst case read operation (similar to a read disturb failure).

4.2.3 Write failure

A write failure event occurs if the value stored in the cell cannot be flipped while the word-line pulse (T_W) is high. The dotted lines shown in Fig. 4.3 shows such a scenario where the cell fails to flip its state within T_W. Due to BTI degradation, the cell transistors can become slower over time and cause write upset. As an example, if we consider that node Q_L in Fig.4.1 stores logic $'0'$ for a long time, the transistors P_R and N_L become slower due to constant BTI stress. If we want to write logic $'1'$, the operation needs to be performed while the word line is high. Otherwise, a write failure occurs. In our analysis, we have considered the worst-case write scenario as defined in [76] and mentioned below.

Worst case write $'0'$ condition:

At t=0, Q_L=0 and Q_R=1.

At t=t_1, write $'1'$ operation is performed at node Q_L ($'0'$ at node Q_R).

At t=t_2, write $'0'$ operation is performed at node Q_L ($'1'$ at node Q_R).

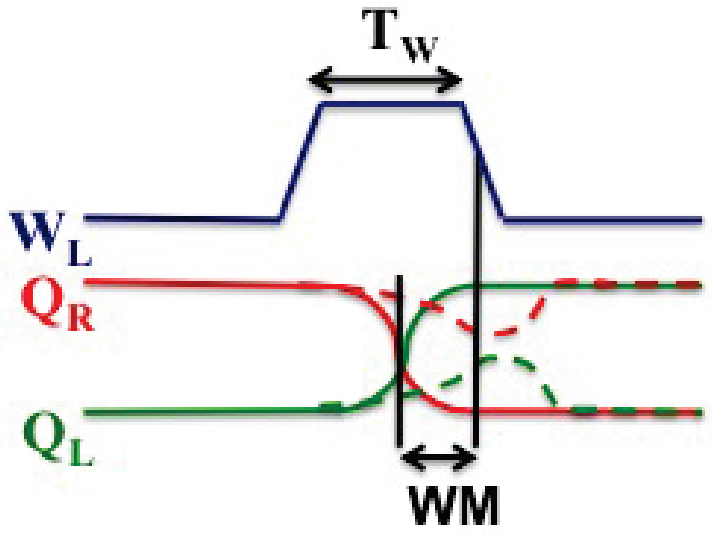


Fig. 4.3. Due to aging effects, $N_R - P_L$ can become weaker, and the cell may fail (dotted lines) to write '0' within T_W. WM is the dynamic write margin when BTI effect is ignored.

Here, the assumptions are: 1) t_1 is long enough to cause considerable BTI degradation to P_R and N_L; and 2) t_2-t_1 is very short and negligible BTI degradation occurs to P_L and N_R.

4.2.4 Data Retention Failure

In standby mode, the V_{DD} of the cell is reduced to a lower value (known as Data Retention Voltage, V_{DR}) in order to reduce the power consumption. In data retention mode, the voltage of the node storing logic $'1'$ is also reduced to a lower value. When the memory cells become active during read or write operations, V_{DR} is again increased to nominal V_{DD}. If the voltage level of the node storing $'1'$ (as an example, Q_L in Fig. 4.1) falls below the trip point of the inverter ($P_R - N_R$), the data stored in the cell can be flipped and lost. This condition is known as data retention or hold failure (Fig.4.4) and can take place if Q_L holds '1' for a long time and transistors P_L and N_R become weak due to BTI effect. The dotted lines correspond to a hold

upset event. Similar to read and write operations, we have considered worst-case hold failure analysis in this Chapter. The condition is similar to the worst case read operation.

4.3 Results and Discussion

In this section, we have investigated the impact of RDF on BTI using worst case read, write, and hold operations of 6T SRAM cells. The failure probability is determined as the fraction of failed SRAM cells due to the mechanisms mentioned above:

$$P_F = \frac{F}{S} \tag{4.2}$$

where, F is the number of failed cells due to read, hold, or write failure mechanisms and S is the total number of samples in our Monte Carlo (MC) simulations. We have considered a memory size of 10KB and unless otherwise specified, a cell ratio (ratio of the widths of pull down to pull up transistors in Fig. 4.1) of 2 is used for our simulations. We have also assumed that the access transistors are not affected by

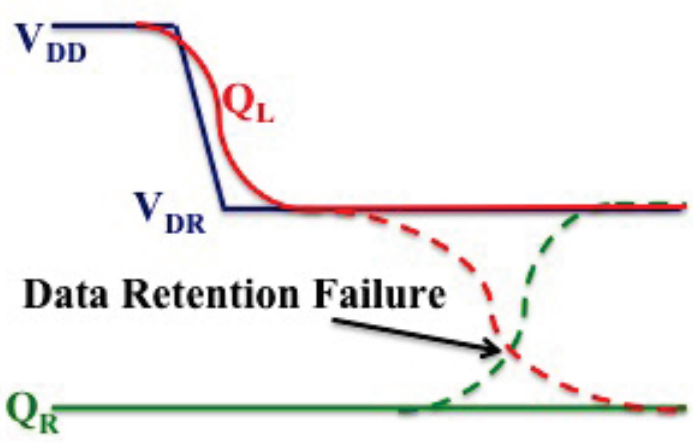


Fig. 4.4. If Q_L stores $'1'$ for a long time, transistors $N_R - P_L$ become slower and the stored data can be flipped and lost (dotted line) when the cell is in standby mode and V_{DD} is reduced to V_{DR}.

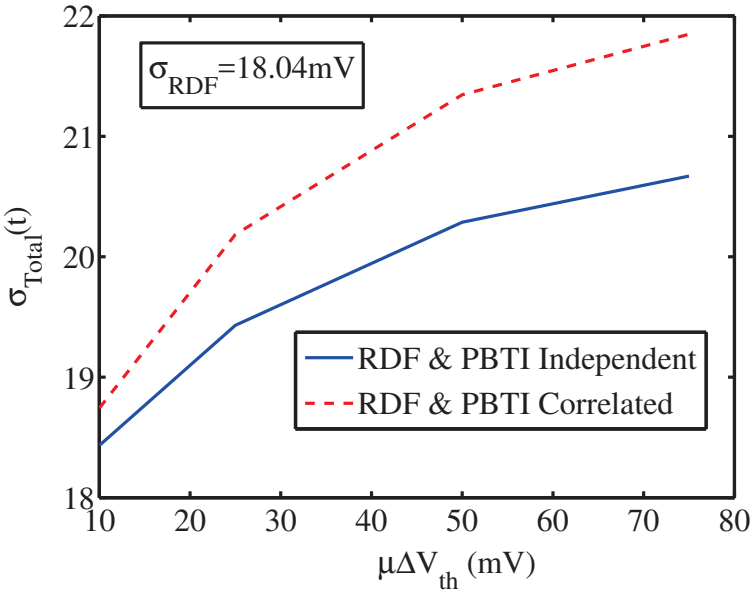


Fig. 4.5. Total V_{th} variability in NMOS transistors with and without considering the impact of RDF on PBTI.

BTI effect since the memory section is relatively inactive and in hold mode for most of the operation [76]. We have also assumed that all transistors are affected by RDF.

Fig. 4.5 shows the total standard deviation (σV_{th}) of V_{th} of NMOS transistors for mean PBTI degradation of 10, 25, 50, and 75 mV with and without considering the correlation between RDF and BTI. We can see that the combined effect is somewhat underestimated if the correlation is ignored. We have used the simulation framework presented in Chapter 3 to determine σV_{th} due to both NBTI and PBTI and the simulation parameters are listed in Tables 2.1 and 3.3.

We have plotted the read access failure probability in Fig. 4.6 with and without considering the correlation between RDF and BTI. T_{Amax} in our simulation is the average cell access time of SRAM cells considering the variability due to RDF. In our simulations, $T_{Amax} \approx 38.4$ ps. We observe that the failure probability is slightly higher if we take the correlation into account. In Fig. 4.7, we have plotted the corresponding failure probabilities for read disturb failure. We have reduced the cell ratio to $\frac{1}{3}$ in

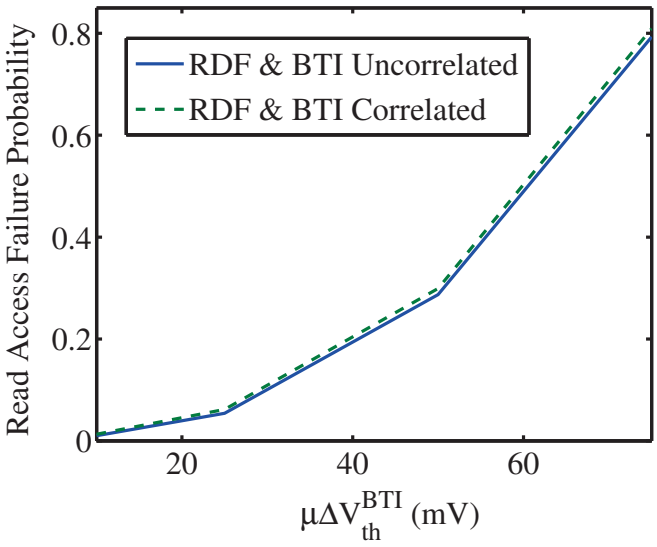


Fig. 4.6. Read access failure probability in 10KB memory array. $T_{Amax} \approx 38.4$ ps in our analysis.

order to observe destructive read operations, even for 10mV mean BTI degradation. The simulation results show the very small effect of correlation on failure probability. In case of write failure analysis, we have set the maximum write access time to the average write time of the cells considering only the RDF variability into account. In our simulations, $T_{Wmax} \approx 142.8$ ps. As shown in Fig. 4.8, the worst write operation is less affected than the read operation, and the impact of the correlation is almost negligible. Finally, we have carried out the worst case hold failure analysis and plotted the results in Fig. 4.9. The data retention voltage is determined as the minimum voltage required to maintain 100% yield, taking only RDF into account. We defined $'$yield$'$ as the percentage of functional bit cells in our analyses. The results show a similar trend to that of the read operation.

We have investigated the relative significance of all failure mechanisms and the importance of the correlation in determining the confidence level of SRAM yield analysis. Fig 4.10(a) shows the difference in yield (ΔYield) in % for all failure mech-

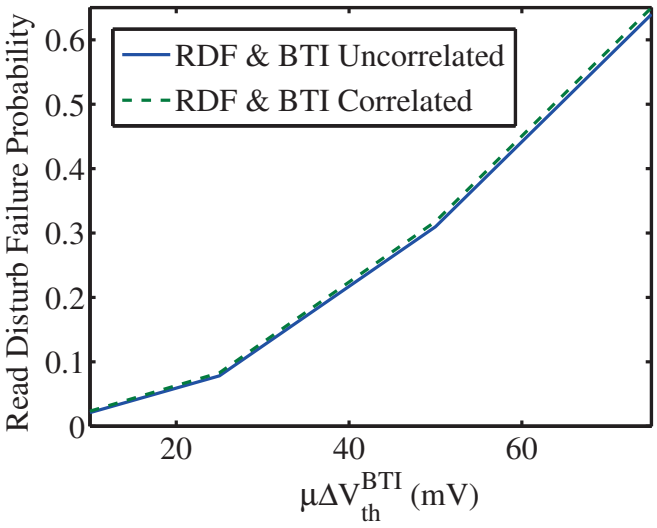


Fig. 4.7. Read disturb failure probability with and without taking the effect of RDF on BTI into account. The failure probability is slightly underestimated when the effect is ignored.

anisms. We see that the impact of RDF on BTI is almost negligible for worst case write operation but somewhat significant in the read and hold failure analyses. The write operation is less affected because the NMOS transistors are wider (stronger) than PMOS transistors in our analysis. In Fig. 4.10(b), we have plotted the number of additional defective cells if the dependance of BTI on RDF is ignored. The total number of faulty cells due to any failure mechanism can be quantified by the following expression [76]:

$$N_D(t) = MemorySize \times P_F(t) \tag{4.3}$$

where, $N_D(t)$ and $P_F(t)$ are the number of defective cells and failure probablity, respectively. The memory size is 10KB in our simulations. We have compared the results with 3σ level (99.7%), and observe that the desired confidence level is not achieved in the case of the Read and Hold failure analyses. Since SRAMs are driven

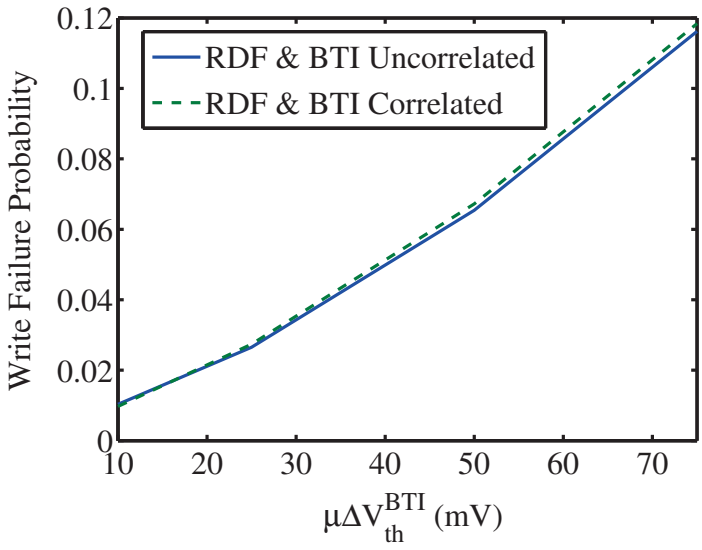


Fig. 4.8. Worst case write failure probability is very small since the pull down transistors are stronger than pull up transistors.

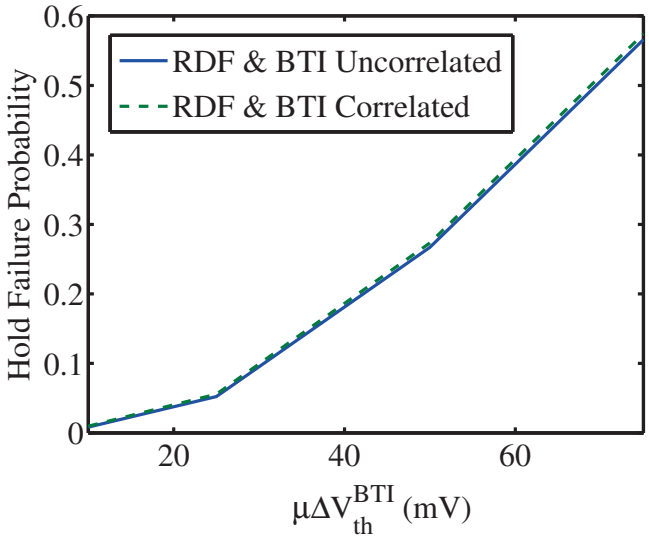
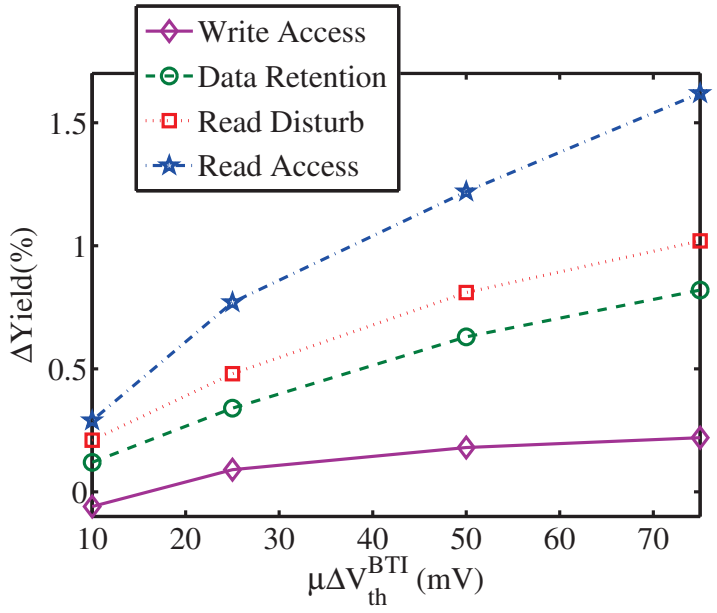
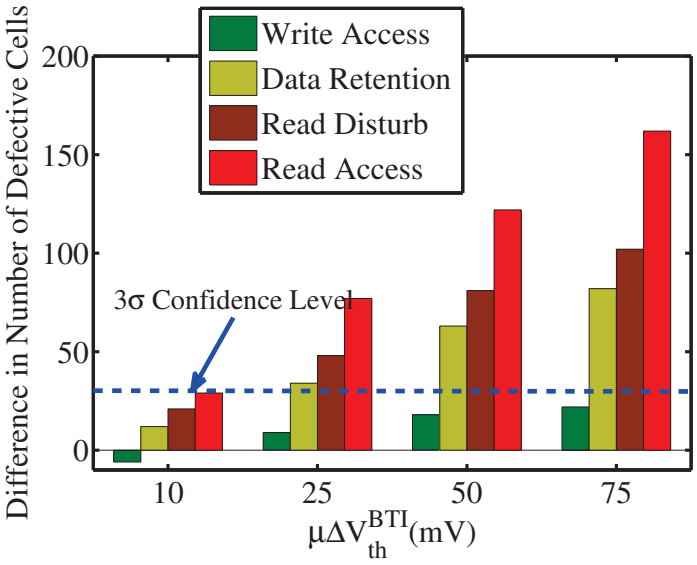


Fig. 4.9. Worst case data retention failure shows very small dependance on the correlation between RDF and BTI. However, this dependance is required to be considered in order to achieve 3σ or higher level of error immunity.



(a)



(b)

Fig. 4.10. (a) The difference in worst case yield is small in the case of read and hold operations and almost negligible in the case of the write operation; (b) The number of faulty cells in the worst case read and hold operations fail to achieve the desired 3σ confidence level when the correlation between RDF and BTI is ignored. We defined 'yield' as the percentage of functional bit cells.

to 6σ confidence level of error immunity (i.e., only 2 faulty cells are acceptable in 1 billion bit cells) [76], this analysis is crucial in accurate yield estimation.

4.4 Conclusion

In this chapter, we have carried out worst case SRAM failure analysis considering the effects of RDF, NBTI, and PBTI. We have observed that the impact is somewhat significant on read and hold failure mechanisms, and can be crucial if we need to achieve a 3 sigma or higher confidence level. Write operation is almost unaffected by this dependence. Therefore, this dependence may have to be considered for accurate yield analysis, and can be significant in error detection and correction techniques of SRAMs.

5. UNIFIED RELIABILITY MODEL OF FINFET TRANSISTORS

The technology roadmap has been evolving in order to keep up with Moore's law. The introduction of High-κ metal gate (HKMG) transistors and FinFET architecture have enabled the scaling down of CMOS technology beyond the 32nm node. However, new failure mechanisms, such as Positive Bias Temperature Instability (PBTI) and Stress Induced leakage current (SILC), have emerged because of this transition. In addition, the corners of tri-gate FinFET transistors can cause 2D diffusion and field crowding that can complicate the reliability assessment of Negative Bias Temperature Instability (NBTI) and Time dependent Dielectric Breakdown (TDDB). In addition, fin thickness plays a role in NBTI degradation. To that effect, we have proposed an unified model of NBTI, PBTI, and TDDB for HKMG FinFET transistors. The simulation framework can be extended to include other reliability issues, such as SILC, Hot Carrier Induced Damage (HCI), and Random Telegraph Noise (RTN). This unified model can be leveraged to 14nm and beyond.

5.1 Introduction

Multi-gate FinFET architecture has replaced conventional planar transistors with their better short-channel effect and higher integration density. In addition to the width and length of the gate, the FinFET structure introduces an additional dimension, the 'fin height'. This third dimension provides enhanced gate control; therefore, the drive current increases and the transistors' off-current decreases without changing the planar area. As a result, better device performance and lower power consumption are achieved [83–85]. However, as the device area is scaled down, the variability and reliability assessment becomes more complicated. Although the aging effects are

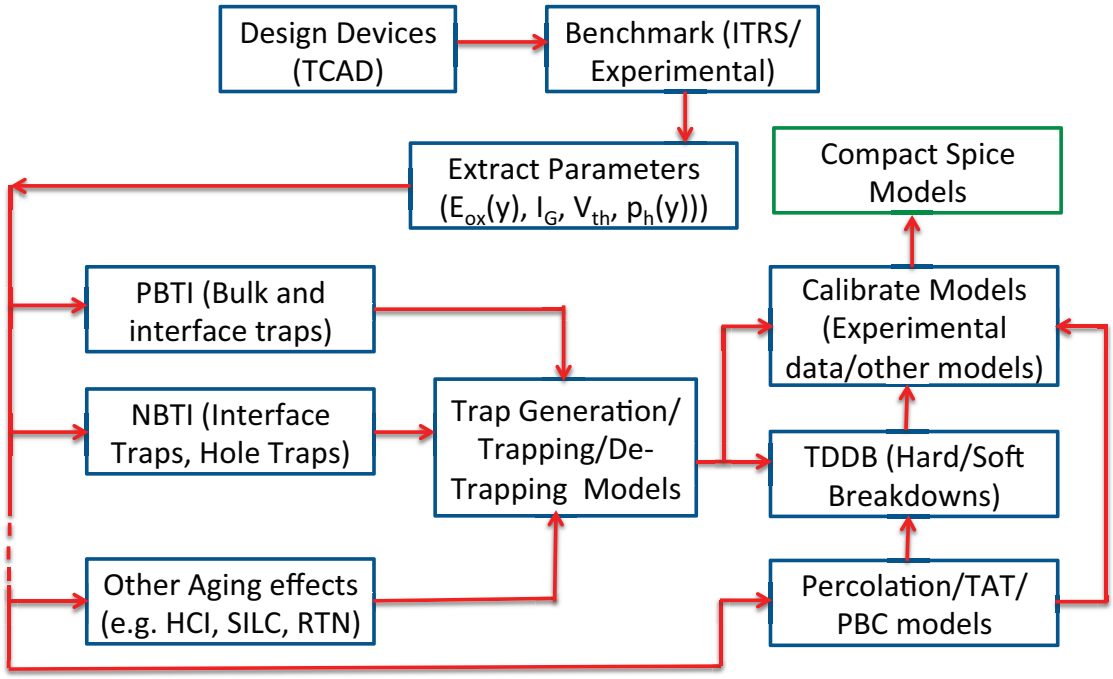


Fig. 5.1. Simulation framework of unified reliability model of HKMG FinFETs.

mostly related to the physics of the high-κ dielectric gate stacks and Si-dielectric interface properties, the corners of the tri-gate architecture can cause faster NBTI and TDDB degradation in the devices. Therefore, a comprehensive study on the reliability analysis of FinFET devices is necessary. In this chapter, we have developed a unified model for BTI and TDDB for FinFET devices and verified the model with experimental results. The models are flexible and the simulation framework can be extended to incorporate other aging mechanisms, such as HCI, RTN, and SILC. In section 5.2, we have explained the simulation framework of this research. In section 5.3, we have discussed the device structure designed in TCAD simulator at 14nm technology node. We have also explored the fin corner effect in this section. Section 5.4 briefly explains the models used in our work. The models are verified with experimental data in section 5.5. We have discussed some of the results in this section. Finally, section 5.6 concludes this Chapter.

5.2 Simulation Framework

Fig. 5.1 shows the simulation framework for FinFET reliability analysis. We have designed our FinFET devices using a TCAD simulator [15]. The device characteristics can be benchmarked with ITRS [86] projection for scaled multi-gate devices or available experimental data. From the device simulator, we can extract oxide electric field (E_{ox}), gate leakage current (I_G), threshold voltage (V_{th}), and inversion channel carrier density (e.g., hole density for PMOS devices) and incorporate them in different aging effect models (BTI, HCI, and SILC). The models need to be calibrated with the experimental data and combined them with the Percolation model for TDDB analysis. Trap Assisted Tunneling (TAT) and Post-Breakdown gate leakage Current (PBC) model can also be integrated into the TDDB model that may require further calibration or validation. From the calibrated device level models, compact Spice models for circuit simulators [45, 87] can be generated, which is part of our future plan.

5.3 Device Structure and Impact of Fin Shape on Performance

We have designed our FinFET devices in TCAD simulator and benchmarked them with ITRS 2013 [86] projection for 14nm technology. Fig. 5.2 shows the 3D view and cross section of our NMOS device. Different design parameters are listed in Table 5.1. In this section, we have also investigated the impact of Fin shape on device performance. The presence of the top gate in a triple-gate structure can cause corner effect [89]. In Fig. 5.3, we have shown the field distribution inside the channel region of a Fin with sharp corners. We can see in Fig. 5.3(b) that the electric field is more concentrated at the corners. We have extracted the field distribution at the corner and at a point inside the channel region corresponding to the mid-point of the top gate along the fin width. The extracted fields are plotted in Fig. 5.4 as a function of the channel length. The field at the corner is approximately $\sqrt{2}$ times higher because two gate planes overlap near the gate corners and increase the effective electric field.

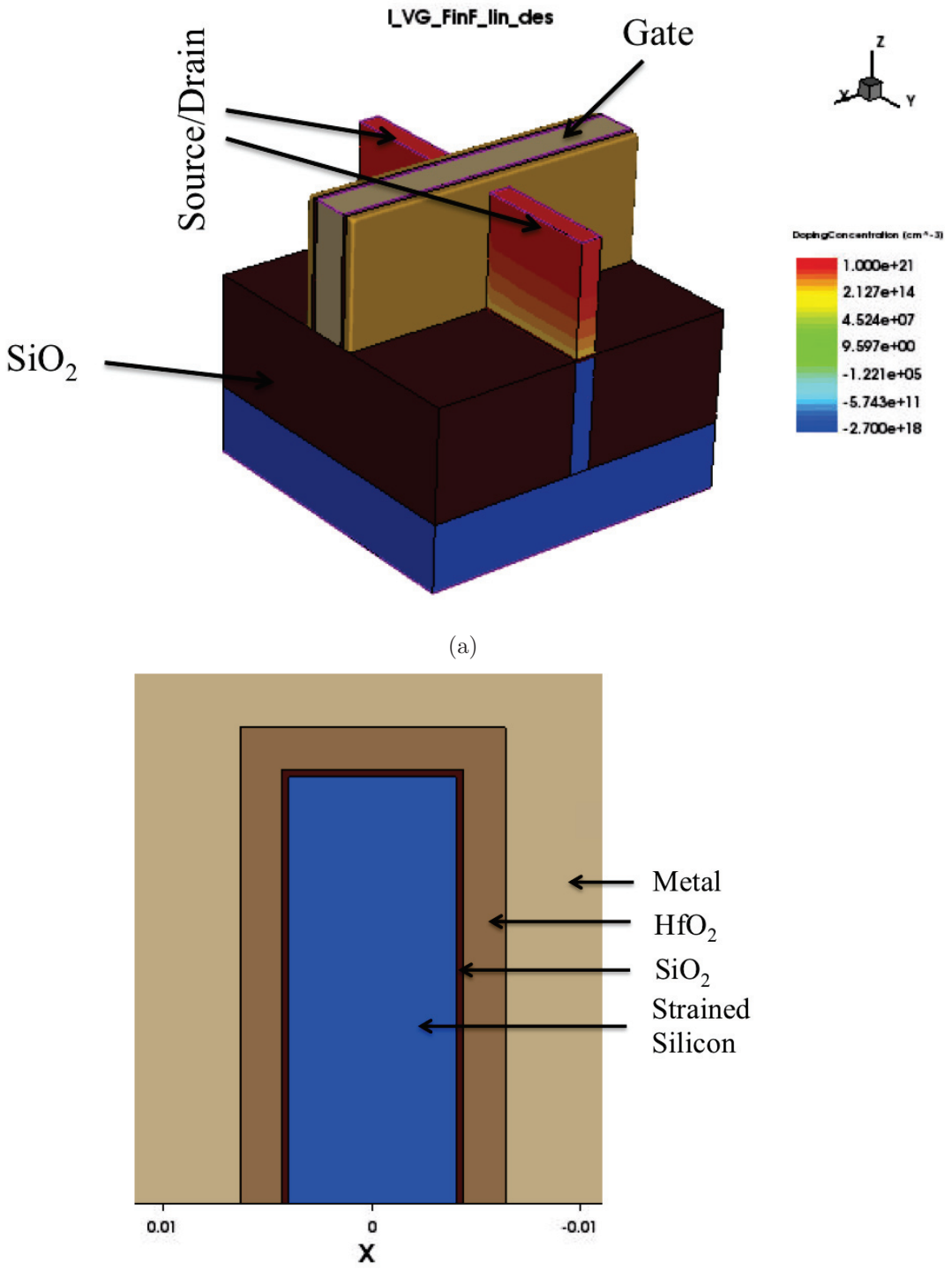
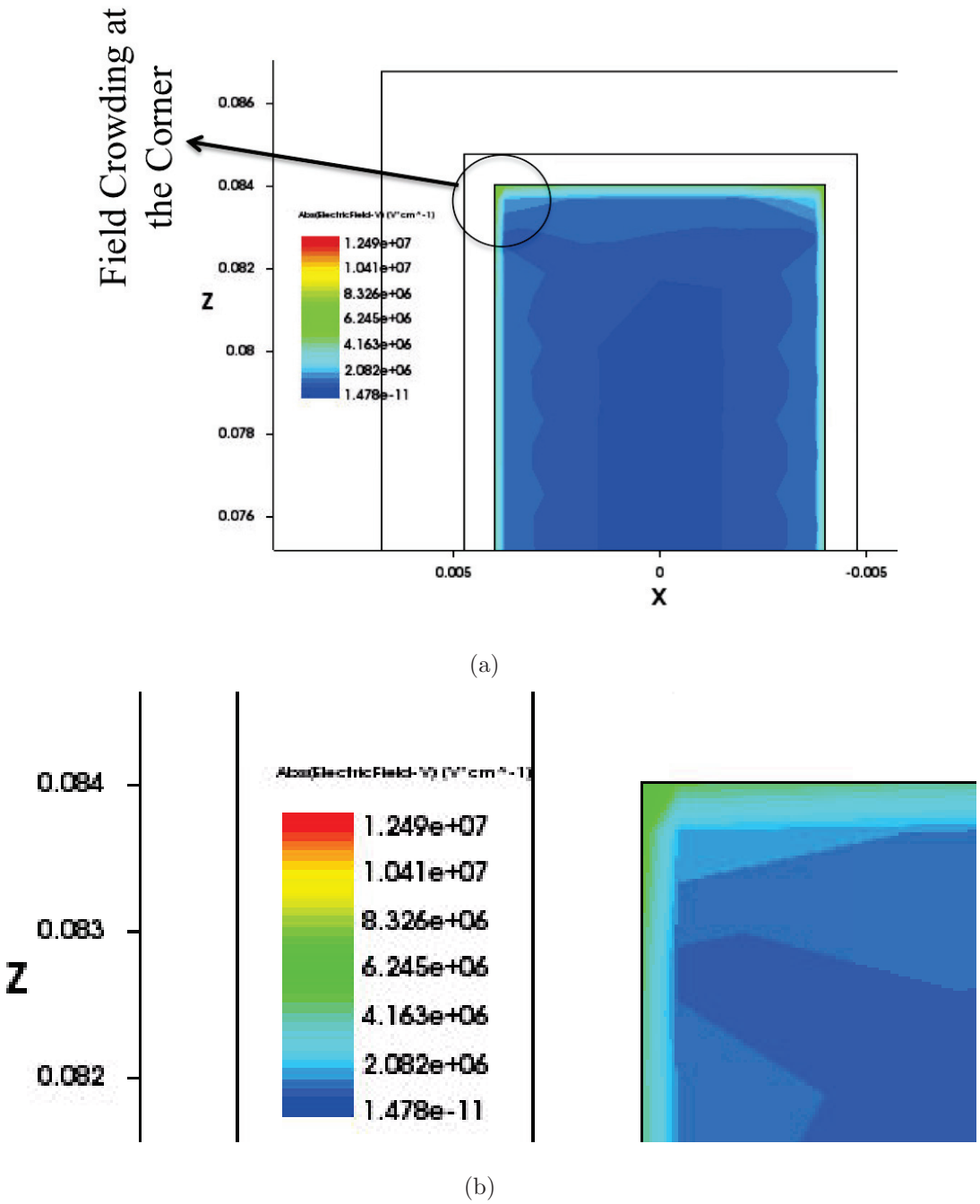


Fig. 5.2. (a) Device Structure of 14nm nFinFET; (b) Cross-section of the device.



(a)

(b)

Fig. 5.3. (a) Electric field in a rectangular sharp Fin; (b) Field crowding at one of the corners.

Table 5.1.
nFinFET Design Parameters

Parameter Name	**Value**	**Reference**
Gate Length	14.4nm	[86]
Supply Voltage	0.85V	[86]
Fin Width	8nm	[88]
Fin Height	42nm	[88]
Gate Oxide Material	HfO_2, SiO_2	
Equivalent Oxide Thickness	≈0.77nm	[86]
Channel Doping	$1x10^{18}/cm^3$	[86]
Source/Drain Doping	$1x10^{20}/cm^3$	Calibration Parameter
I_{Dsat}	1680 $\mu A/\mu m$	[86]
I_{off}	100 nA/μm	[86]
NMOS Gate Workfunction	4.42eV	Calibration Parameter

This field crowding can potentially cause faster degradation of the oxide layers at the corners and reduce the overall lifetime of the devices. However, this effect can be eliminated by adding a rounded shape to the corners. We have shown the electric field distribution of a FinFET with rounded fin corners in Fig. 5.5. In Fig. 5.6, we have shown a similar plot with a rounded trapezoidal fin shape, which is more realistic in present CMOS technology. We observe in both cases that the rounded shape can completely eliminate the field crowding effect. In order to compare the nominal performance of devices with different fin shapes, we have plotted their $I_G - V_G$ characteristics in Fig. 5.7 for $V_D = V_{DD}$. All the three transistors were designed with equal fin area and on-current. We observe in Fig. 5.7 (b) that the sub-threshold leakage is slightly higher for trapezoidal FinFETs. A possible reason could be the widening of fin thickness near the bottom of the trapezoidal structure. This may have caused relatively worse gate control and higher leakage current. For our reliability analysis, we have used rectangular FinFETs with rounded corners.

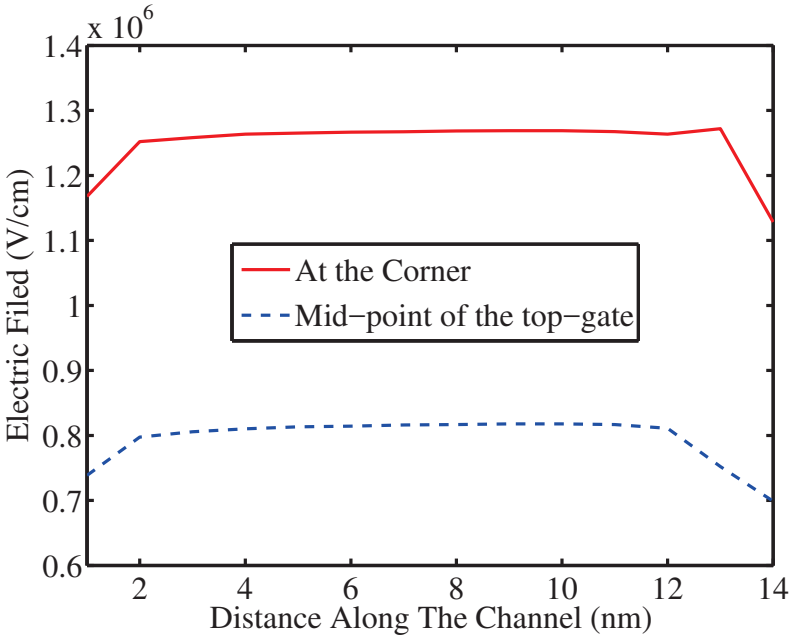


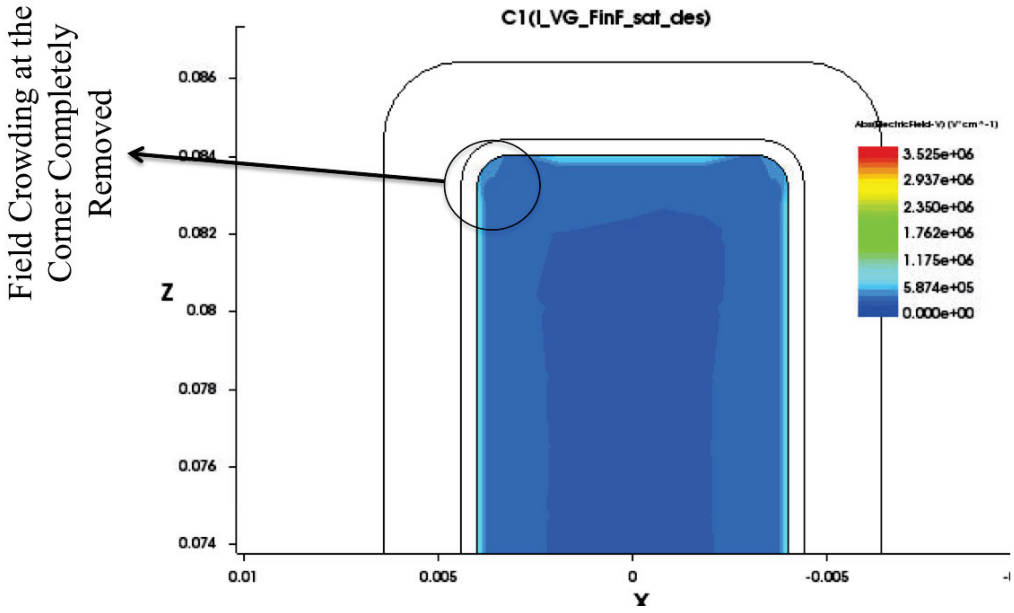
Fig. 5.4. Comparison of Electric field in the channel region between one of the corners and at a point corresponding to the mid-point of the top gate.

5.4 BTI and TDDB in FinFET transistors

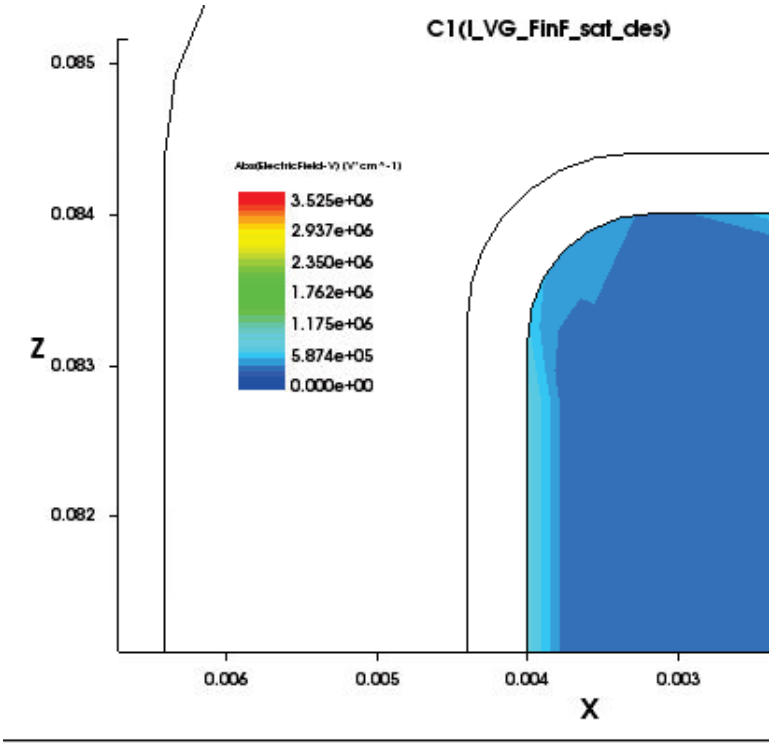
In this section, we have briefly explained the models we implemented for FinFET architecture.

5.4.1 Positive Bias Temperature Instability

PBTI is a mechanism related to the quality of HK materials, particularly in NMOS transistors [9]. Therefore, the statistical model developed in Chapter 2 is equally applicable for FinFETs. Both pre-existing and stress-induced defects are responsible for PBTI degradation. Since the pre-existing traps are fast saturated [11], their contribution is almost constant to a device under constant stress. On the other hand, the stress-induced defects accumulate in the dielectric layers and cause gradual degrada-

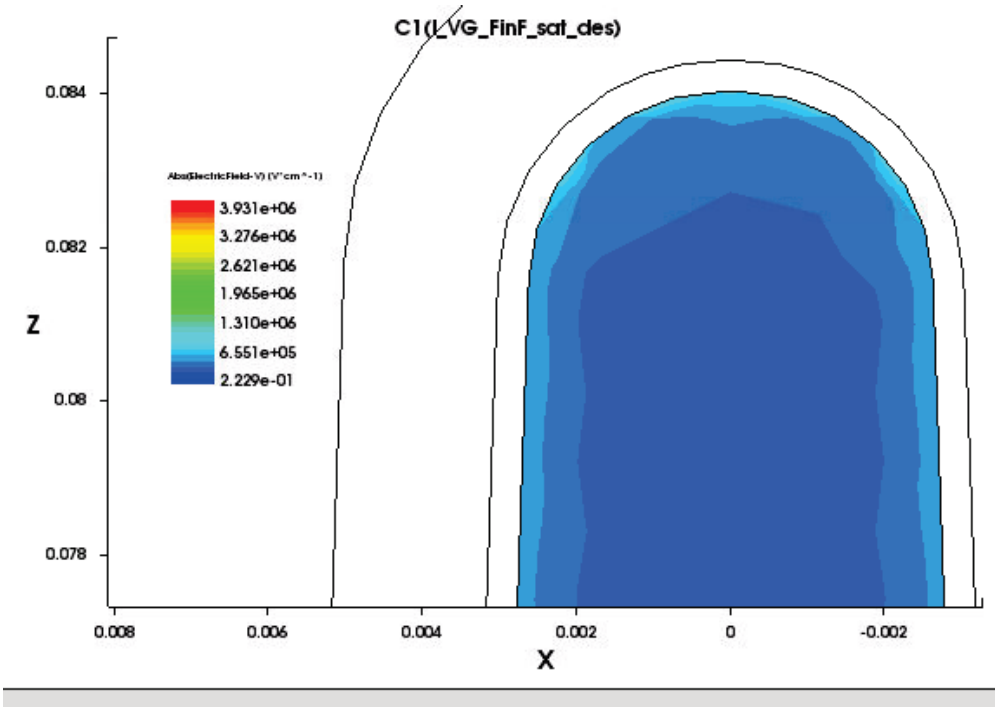


(a)

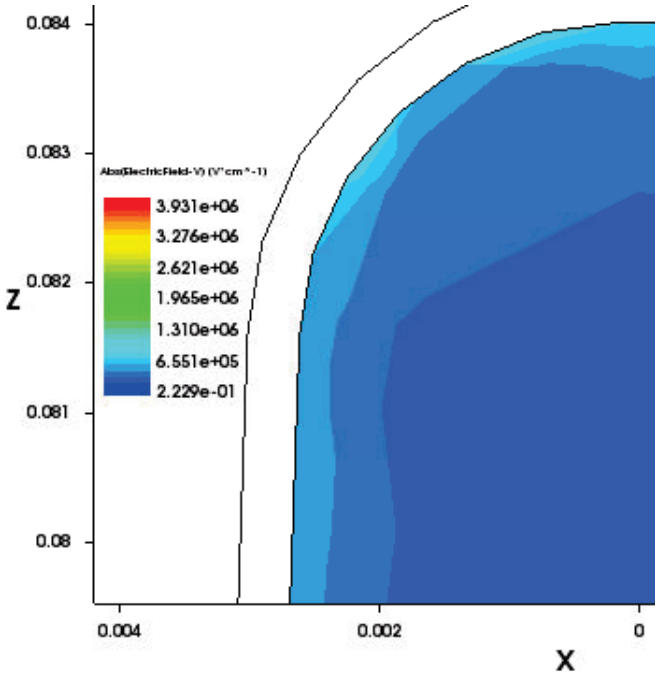


(b)

Fig. 5.5. (a) Electric field in a rectangular rounded corner Fin; (b) Field crowding is eliminated at the rounded corners.

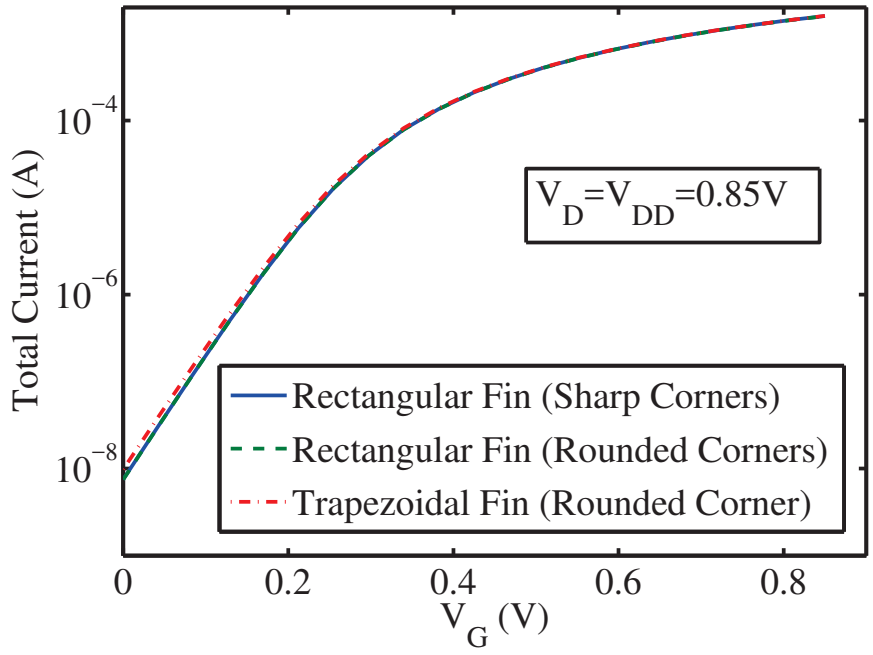


(a)

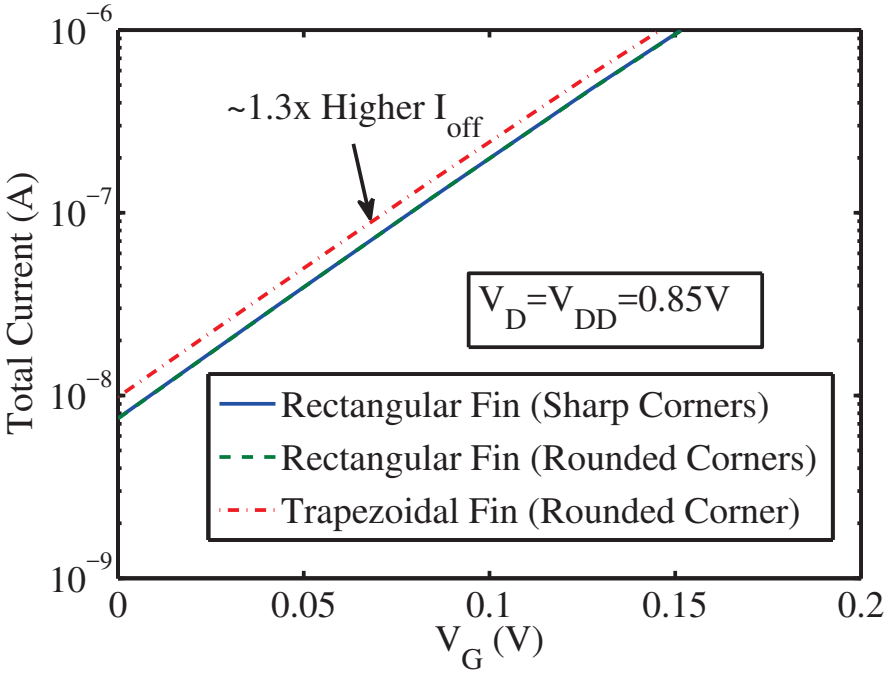


(b)

Fig. 5.6. (a) Electric field in a trapezoidal rounded Fin; (b) Field distribution inside the channel region.



(a)



(b)

Fig. 5.7. Performance comparison of FinFETs with different Fin shapes: (a) $I_D - V_G$ Comparison; (b) I_D current in the sub-threshold region.

tion to the device. The dependance of trap generation on stress period and voltage can be expressed as power laws. In addition, this stress temperature dependance can be modeled with the Arrhenius relationship. The combined effect of stress period, voltage, and temperature on generated traps N_{Gi} can be expressed using the following equation [49]:

$$N_{Gi}(V_G, T, t) = A(V_{Gi})^{\gamma} exp(-E_a/K_B T)t^n \tag{5.1}$$

where, n is the PBTI power exponent, γ is the voltage acceleration factor, E_a is the temperature activation energy, and $'A'$ is the proportionality constant and calibration parameter. By applying probability theory, we can determine the total number of charged traps, $N_C(t)$ for a certain stress condition V_G, T, t using (5.2) [49]:

$$N_C(t) = \int_{x=1}^{N_W} \int_{y=1}^{N_L} \int_{z=1}^{N_{T_{di}}} P(C)_{t,x,y,z} \mathrm{d}x\mathrm{d}y\mathrm{d}z \tag{5.2}$$

where, N_w, N_L, and N_{Tdi} are the number of discrete points considered along the width, length, and thickness of the i_{th} dielectric layer, respectively. $P(C)_{t,x,y,z}$ in (5.2) indicates the charging probability of a trap located at (x, y, z) at a certain time, t. The detail of this model, presented in Chapter 2, will not be covered here. ΔV_{th} can be determined using Poisson's equation (Chapter 2). As an alternative approach, the TCAD simulator can also be used for this purpose (Chapter 3). The PBTI simulation parameters used in this chapter are listed in Table 2.1.

5.4.2 Negative Bias Temperature Instability

NBTI is a phenomenon mostly related to the Si-SiO_2 interface trap generation or pre-existing hole traps inside the oxide layer [70]. The 1D Reaction Diffusion (R-D) model is widely used for planar devices in order to model the interface trap generation due to NBTI stress. We have briefly explained the R-D model in Chapter 3. In the case of a triple-gate structure, the corners can introduce 2D diffusion (Fig. 5.8) and

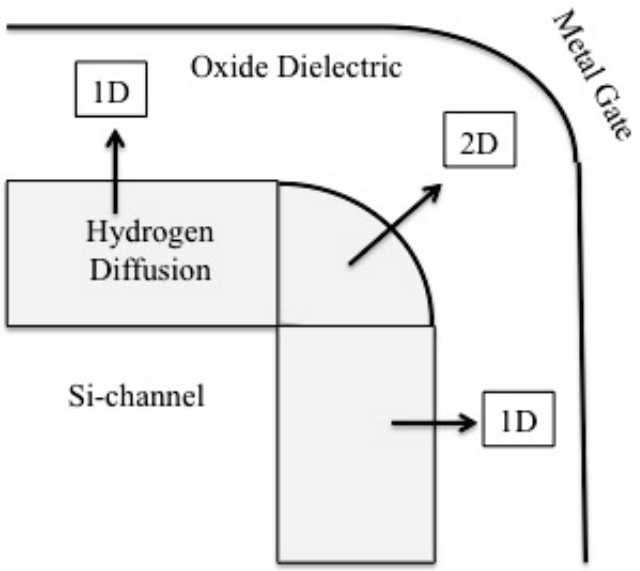


Fig. 5.8. 2D Diffusion at the corners of tri-gate FinFETs [90]

cause faster degradation to the device [90]. In addition, the fin width also plays a role in NBTI degradation. The energy band bending and hole concentration increase in narrow Fins are due to the injection and accumulation of majority carrier electrons in the Fin body [91, 92]. Therefore, it is necessary to incorporate both effects in the existing models in order to accurately predict NBTI degradation. In our extended R-D model, we have used both 1D and 2D diffusion, depending on the location of the traps:

$$N_{IT}(t) = \begin{cases} N_{2D}(t) & \text{At the Corners} \\ N_{1D}(t) & \text{Otherwise} \end{cases}$$

The number of interface traps due to 1D diffusion can determined by [70]:

$$N_{1D}(t) \propto (\frac{k_f N_0}{2k_r})^{2/3}(DH_2 t)^n \tag{5.3}$$

where, k_f, k_r, N_0, and DH_2 are Si-H bond breaking rate, Si-H bond annealing rate, pre-stress Si-H bond density at the Si/SiO_2 interface, and the diffusion coefficient of H_2. In case of 2D diffusion, we need to solve the 2D diffusion equation in the cylindrical coordinate system [90]. The final expression is given below:

$$N_{2D}(t) \propto (\frac{\pi}{12})^{1/3}(\frac{k_f N_0}{k_r})^{2/3}(DH_2t)^{2n} \tag{5.4}$$

The derivation of equation 5.4 is shown in the Appendix A. The parameter k_f depends on the hole density inside the channel region (p_h), oxide electric field (E_{ox}), and stress temperature (T):

$$k_f^{T_{Fin}}(y) = k_{f0}p_h(y)exp(\gamma E_{ox}(y))exp(-\frac{E_{AF}}{K_B T}) \tag{5.5}$$

where, γ is the field acceleration factor, E_{AF} is the forward activation energy, and K_B is the Boltzmann constant. k_{f0} in (5.5) is a proportionality constant and the calibration parameter in our model. E_{ox} and p_h inside the channel region are functions of the channel length y [93]. In addition, p_h depends on fin thickness. In order to take these effects into consideration, we have designed our pfet device in the TCAD simulator and extracted both E_{ox} and p_h as a function of the channel length for a given fin width (T_{Fin}). The NBTI simulation parameters used in this Chapter are listed in Table 3.3.

5.4.3 Time Dependent Dielectric Breakdown

The Equivalent Oxide Thickness (EOT) is very thin (less than 0.8nm) in technology nodes of 14nm and smaller [86]. In addition, unless taken care of in the device design, the corner effect can cause faster degradation of tri-gate FinFETs. Therefore, TDDB analysis is crucial to predict the lifetime of a device. The trap generation models developed for NBTI and PBTI can be incorporated into the percolation model in order to estimate the TDDB lifetime of the devices. In case of nFinFETs, the same PBTI traps contribute to TDDB; in the case of pFinFETs, bulk oxide traps as well as NBTI interface traps play a role in forming gate-body percolation paths [31]. Therefore, both NBTI and PBTI models are necessary for TDDB analysis. We have integrated the NBTI and PBTI trap generation models with the 3D percolation model

implemented in Chapter 3. The analytical expression of Weibull slope (β) for the 1^{st} Soft BreakDown (SBD) can be determined from the following expression:

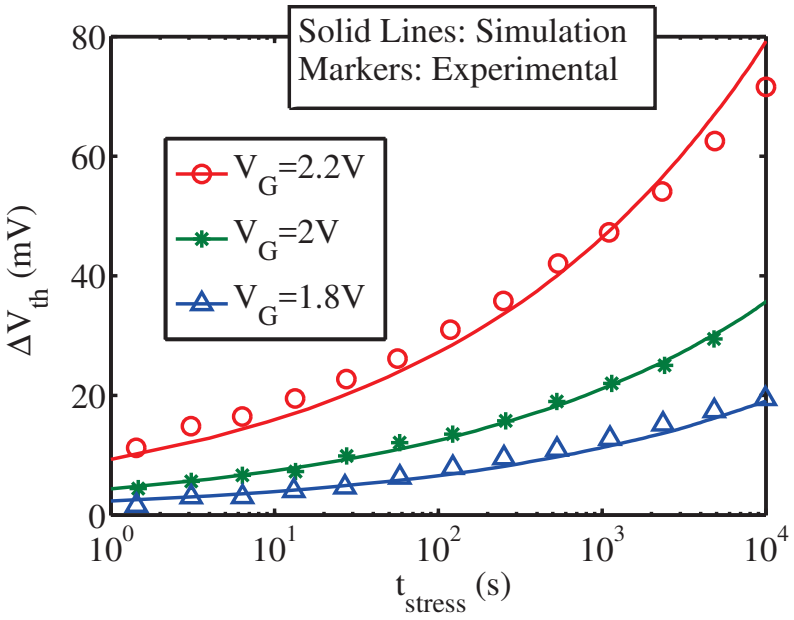
$$\beta = \alpha(\frac{T_{ox}}{a_0}) \tag{5.6}$$

where, α is the time exponent for defect generation, a_0 is the cell or defect size, and T_{ox} is the thickness of the HK or IL layer.

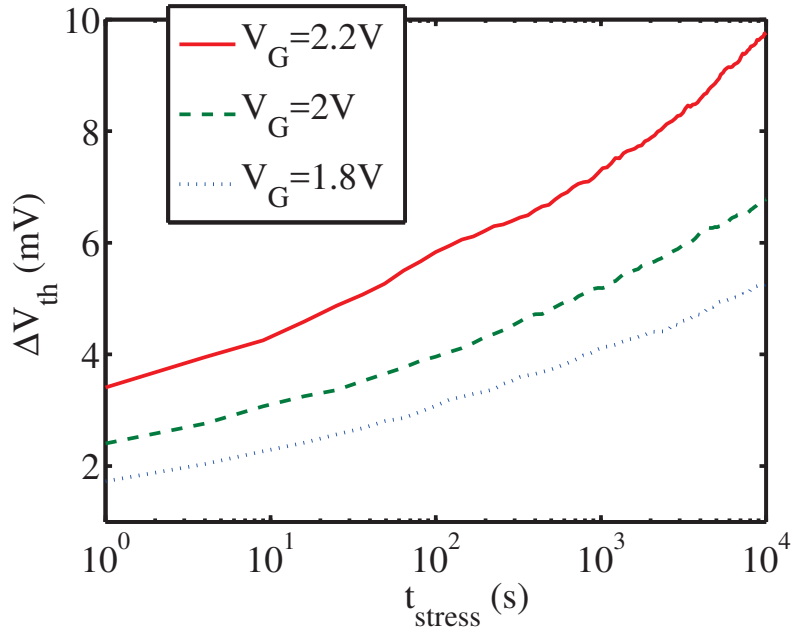
5.5 Results and Discussion

We have verified our PBTI model by the experimental results presented in [92] for an equivalent oxide thickness of 1.1 nm. We have calibrated the proportionality constant $'A'$ in (5.1) to match the average degradation of 1000 sample devices with experimental data [92] for three different stress voltages (2.2V, 2V, and 1.8V) and plotted the results in Fig. 5.9(a). The stress temperature is maintained at 125^oC. The standard deviations for the corresponding stress conditions are shown in Fig. 5.9(b). In Fig. 5.10, we have shown similar plots for NBTI degradation in pFinFETs. Stress voltages of -2.2V, -2V, and -1.8V were applied at 125^oC, and the average degradation of 1000 devices are matched with experimental results [92]. For both n and p type devices, the standard deviation follows a power law dependence over time with a power exponent of $n/2$, where n is the PBTI or NBTI power exponent. In Fig. 5.11, we have plotted the standard deviation as a percentage of corresponding average values for both N and PBTI (on left and right y axes, respectively). Variability due to BTI trap generation is higher in the case of NBTI. The above observations are in agreement with what we have observed for planar devices in Chapter 2.

Next, we calibrated our TDDB model using experimental data of nFinFETs [94] for three different stress voltages (2.2V, 2.3V, and 2.4V). The PBTI trap generation model is calibrated to match the Weibull distribution for HBDs. The Weibull slope β reported for SBDs in 0.8nm EOT devices is 1.08 [94]. The authors in [94] claimed that the slope indicates roughly three defects in a SBD path. Since, $\beta_{HBD} = k * \beta_{SBD}$, where, k is the number of soft breakdowns required to cause a hard breakdown.

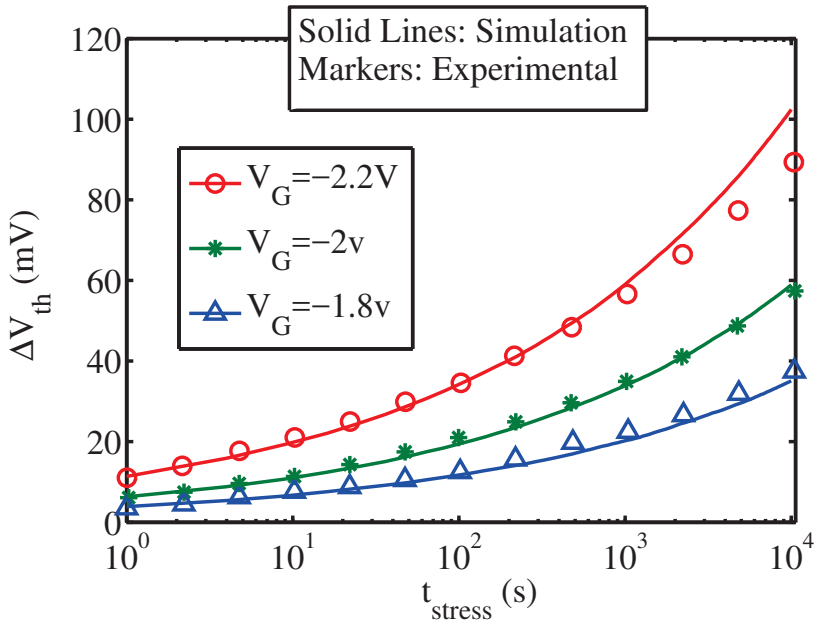


(a)

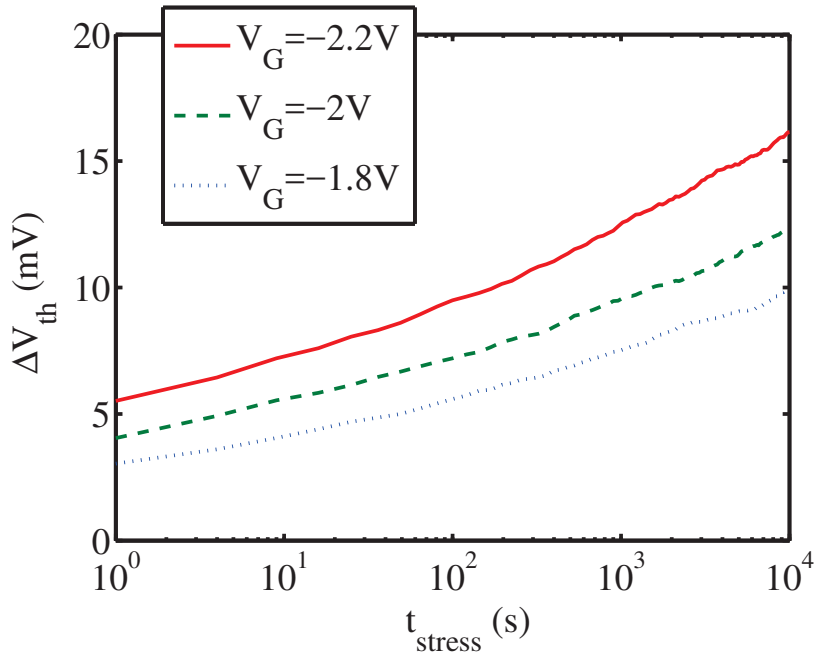


(b)

Fig. 5.9. ΔV_{th} degradation due to PBTI: (a)Mean degradation; (b) Standard deviation. PBTI power exponent, n_{PBTI} ≈0.23.



(a)



(b)

Fig. 5.10. ΔV_{th} degradation due to NBTI: (a)Mean degradation; (b) Standard deviation.NBTI power exponent, n_{NBTI} ≈0.24.

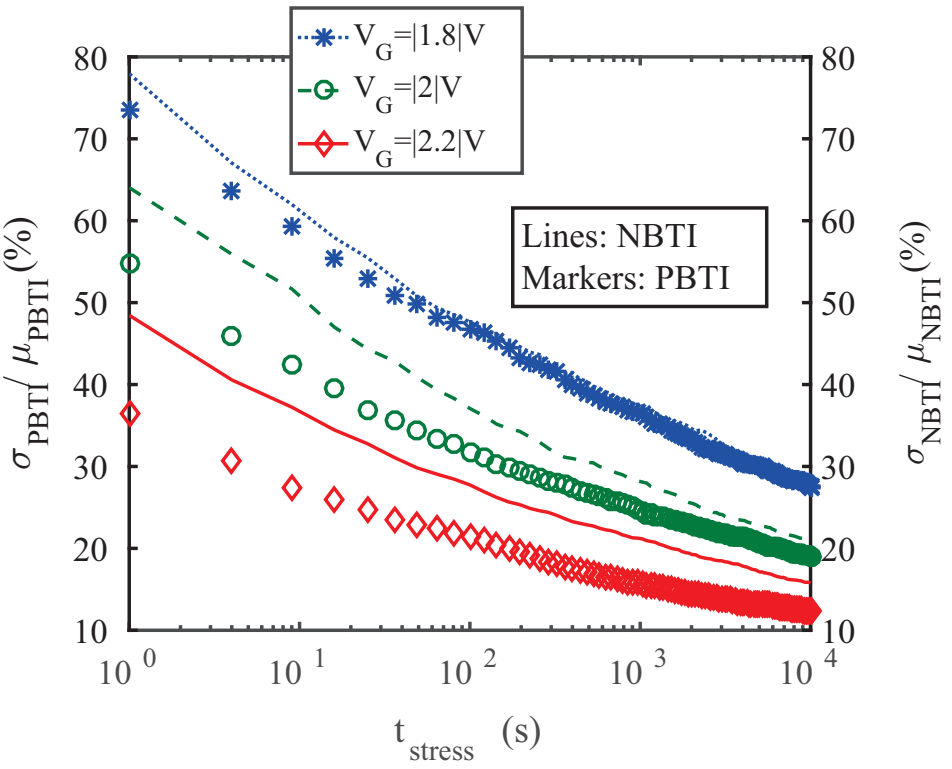


Fig. 5.11. Comparison of NBTI and PBTI variability. Similar to the planar devices, degradation due to NBTI is higher than PBTI.

We determined the trap generation power exponent α ≈0.36 for these devices and assumed same trap generation rate in both HK and IL layers. We extracted the Weibull distribution data from [94] and determined that the slope is ∼1.8. This indicates that roughly two soft breakdowns have caused hard breakdown in those devices. We calibrated the trap generation model using the parameters listed in Table 2.1 and plotted the Weibull distribution for the second soft breakdown. The area of our simulated devices is 1324 nm^2 ($= A_1$), which is different than the devices (A_2 =3.7x10^{-9} cm^2) used in [94]. Therefore, we have applied area scaling to our Weibull distribution using the following expression [31]:

$$W_2 = W_1 + ln(\frac{A_2}{A_1}) \tag{5.7}$$

where, W_1 and W_2 are the Weibull distribution corresponding to device areas A_1 and A_2, respectively. We have run Monte Carlo (MC) simulations for 10000 devices. The simulation results are in good agreement with the experimental data.

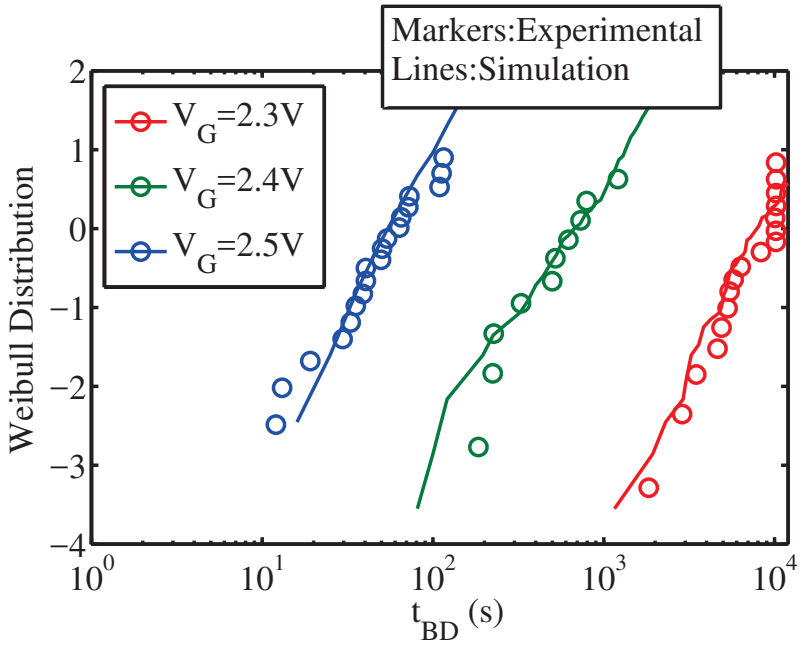


Fig. 5.12. The trap generation model of PBTI is calibrated to match the time to breakdown (t_{BD}) due toTDDB is calibrated with experimental results.

5.6 Conclusion

In this chapter, we have focused on BTI and TDDB in scaled CMOS technology, particularly in FinFETs. We have investigated the impact of tri-gate corner effects on device performance. We have also calibrated the models with the experimental data reported for FinFET devices. The objective of this work is to develop a unified-trap generation model that can be calibrated for BTI degradation and leveraged to TDDB analysis using the same trap generation models. The simulation framework can be extended to include other aging effects, such as HCI, RTN, and SILC. In addition, TDDB gate leakage current models can also be integrated to the simulation framework. The models can be leveraged to generate compact models for circuit simulation, which is part of our future plan.

6. CONCLUSION AND FUTURE DIRECTION

The scaling down of technology leads to faster switching speed and larger integration density. However, these benefits come at the expense of increased complexity in reliability assessment. As a result, the flexibility in allowed margins for process as well as time-dependent variability decreases in every technology generation. In order to combat reliability challenges, a conventional VLSI design incorporates a certain guard band so the circuits can sustain about 10-15% of parametric variations over the projected lifetime and bypass any possible correlation between process-induced variability and time-dependent reliability issues [69]. However, as we go down deeper into the nanometer technology (45nm and beyond), the design criteria become more and more strict. In addition, correlation between time-zero and aging effects can play important role in the performance assessment of memory devices. Hence, any possible correlation among randomly distributed parameters should not be ignored. To that effect, we proposed a statistical framework that takes into account the correlation between time-zero and time-independent variability.

In Chapter 2, we proposed a stochastic model for parametric V_{th} degradation due to oxide trap generation. The model is scalable and can be calibrated to different sets of experimental results. In Chapter 3, we extended our model to analyze time-dependent oxide breakdowns. We designed HKMG transistors with a gate length of 30nm in Sentaurus device simulator. We integrated our PBTI model with the device simulator and analyzed the correlation between discrete dopant-induced variability and PBTI. We have also integrated the WKB approximation-based SILC analysis and 3D percolation model with the PBTI trap generation model and developed a unified model for PBTI, TDDB, and SILC analyses. We concluded that the V_{th} variation worsens due to this correlation. However, TDDB and SILC estimations remain almost unaffected. We have also integrated a statistical NBTI model to our

simulation framework in order to carry out circuit analysis. The simulation results show that the device level behavior is translated to circuit behavior; both follow similar trends in performance degradation. Apart from this, the impact of correlation appears to lessen when circuit complexity increases.

Coexistence of process induced variation and time dependent parametric variation and their correlation can make the design of a robust memory device more challenging. In Chapter 4, we have leveraged our simulation methodology to analyze SRAM transient behavior and failure mechanisms. We observe that a moderate number of bit cells fail to meet the 3σ confidence level in case of read and hold analyses. Hence, this correlation needs to be considered in the reliability quantification of circuits.

In Chapter 5, we have extended our models developed in Chapter 2 and 3 to FinFET technology. We have investigated the effect of field crowding in triple-gate FinFET devices that can cause faster degradation of the oxides in the corner regions. We have observed that corner rounding can completely eliminate the effect. In addition, we have incorporated the 2D diffusion of H_2 and the effect of narrow Fins in the conventional Reaction-Diffusion model of NBTI degradation. Since PBTI degradation is related to the physics of HK materials, the model developed in Chapter 2 is applicable for FinFET technology without the need for any further modification. We have incorporated the 3D statistical percolation model to our BTI models in order to estimate the lifetime of FinFET devices. The models are verified with the published experimental data.

6.1 Future Work

The models and simulation framework developed in this research can be leveraged to integrate other aging effects. In addition, the circuit behavior is even more crucial than the individual devices. Researchers have shown that a circuit can still be functional even after some of the individual devices within the circuit completely fail [95]. Therefore, it is important to develop compact models for different aging effects that

can reasonably predict the circuit performance under the corresponding aging effects. Therefore, we can continue this research work with the following goals.

6.1.1 Extend the models and integrate other aging effects

SILC analysis was part of our unified model for planar devices developed in Chapter 3. We can integrate the Trap Assisted Tunneling current model to the simulation framework for FinFETs. Moreover, RTN and HCI are two other reliability concerns for MOS devices. HCI is related to both interface and bulk trap generations that can cause permanent damage while RTN occurs mostly due to trapping/de-trapping of defects. Our models developed for PBTI and TDDB, and the modified R-D model developed for NBTI analysis in FinFETs, can be leveraged to HCI and RTN analysis. In addition, the post break down leakage current model is necessary to analyze circuit performance and behaviors after the breakdown of individual devices. The models can be incorporated in our simulation framework, as shown in Fig. 1.1.

6.1.2 Compact models for circuit Analysis

We would like to be able to carry out a reliability analysis of Integrated Circuits (IC) with our unified reliability models developed for FinFET devices. To that effect, we plan to develop compact Spice models that can be integrated with commercial circuit simulators [45, 87]. With our unified models, we would also like to carry out the decoupling and rapid acceleration of different aging effects at the IC level.

REFERENCES

REFERENCES

[1] L. Yiming, S. M. Yu, J. R. Hwang, and F. L. Yang, "Discrete Dopant Fluctuations in 20-nm/15-nm-Gate Planar CMOS," *IEEE Transactions on Electron Devices*, vol. 55, no. 6, pp. 1449–1455, 2008.

[2] Y. Li, C. H. Hwang, T. Y. Li, and M. H. Han, "Process-Variation Effect, Metal-Gate Work-Function Fluctuation, and Random-Dopant Fluctuation in Emerging CMOS Technologies," *IEEE Transactions on Electron Devices*, vol. 57, no. 2, pp. 437–447, Feb 2010.

[3] A. Asenov and S. Saini, "Suppression of random dopant-induced threshold voltage fluctuations in sub-0.1-μm MOSFET's with epitaxial and δ-doped channels," *IEEE Transactions on Electron Devices*, vol. 46, no. 8, pp. 1718–1724, 1999.

[4] A. Asenov, A. Brown, J. Davies, S. Kaya, and G. Slavcheva, "Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETs," *IEEE Transactions on Electron Devices*, vol. 50, no. 9, pp. 1837–1852, 2003.

[5] M. A. Alam, "ECE 695A Reliability Physics of Nanotransistors," Jan 2013. [Online]. Available: https://nanohub.org/resources/16560

[6] J. Hicks, D. Bergstrom, M. Hattendorf, J. Jopling, J. Maiz, S. Pae, C. Prasad, and J. Wiedemer, "45nm Transistor Reliability," *Intel Technology Journal*, vol. 12, no. 2, pp. 131–144, 2008.

[7] *International Technology Roadmap for Semiconductors*, 2009. [Online]. Available: http://www.itrs2.net/itrs-reports.html

[8] A. Kerber and E. Cartier, "Reliability Challenges for CMOS Technology Qualifications With Hafnium Oxide/Titanium Nitride Gate Stacks," *IEEE Transactions on Device and Materials Reliability*, vol. 9, no. 2, pp. 147–162, 2009.

[9] S. Ramey and et al., "Intrinsic transistor reliability improvements from 22nm tri-gate technology," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, April 2013, pp. 4C.5.1–4C.5.5.

[10] J. Yang, M. Masuduzzman, J. Kang, and M. Alam, "SILC-based reassignment of trapping and trap generation regimes of positive bias temperature instability," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, 2011, pp. 3A.3.1–3A.3.6.

[11] S. Mahapatra, V. Maheta, A. Islam, and M. Alam, "Isolation of NBTI Stress Generated Interface Trap and Hole-Trapping Components in PNO p-MOSFETs," *IEEE Transactions on Electron Devices*, vol. 56, no. 2, pp. 236–242, 2009.

[12] A. Kerber and T. Nigam, "Challenges in the characterization and modeling of BTI induced variability in metal gate / High-κ CMOS technologies," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, 2013, pp. 2D.4.1 – 2D.4.6.

[13] S. Zafar, A. Callegari, E. Gusev, and M. Fischetti, "Charge trapping related threshold voltage instabilities in high permittivity gate dielectric stacks," *Journal of Applied Physics*, vol. 93, no. 11, pp. 9298–9303, 2003.

[14] J. Yang, M. Masuduzzaman, K. Joshi, S. Mukhopadhyay, J. Kang, S. Mahapatra, and M. Alam, "Intrinsic correlation between PBTI and TDDB degradations in nMOS HK/MG dielectrics," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, 2012, pp. 5D.4.1–5D.4.7.

[15] *Sentaurus Device Simulator*, G-2012.06 ed. Mountain View, California: Synopsys, Inc., 2012. [Online]. Available: http://www.synopsys.com/Tools/silicon/tcad/

[16] H. Y. Yang, H. Niimi, and G. Lucovsky, "Tunneling currents through ultrathin oxide/nitride dual layer gate dielectrics for advanced microelectronic devices," *Journal of Applied Physics*, vol. 83, no. 4, pp. 2327–2337, 1998.

[17] A. Goud, *Asymmetric Underlap Optimization of Sub-10nm FinFETs for Realizing Energy-Efficient Logic and Robust Memories*. Ph.D. Dissertation, Purdue University, West Lafayette, IN, USA, 2016.

[18] K. Mistry and et al, "A 45nm Logic Technology with High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging," in *Proc. IEEE International Electron Devices Meeting (IEDM)*, 2007, pp. 247–250.

[19] X. Chen and et al, "A cost effective 32nm high-K/ metal gate CMOS technology for low power applications with single-metal/gate-first process," in *Proc. Symposium on VLSI Technology*, 2008, pp. 88–89.

[20] N.-H. Hsu, J. W. You, H. C. Ma, S. C. Lee, E. Chen, L. S. Huang, Y. C. Cheng, O. Cheng, and I. C. Chen, "Intrinsic hot-carrier degradation of nMOSFETs by decoupling PBTI component in 28nm high-κ/metal gate stacks," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, 2012, pp. XT.13.1–XT.13.4.

[21] J. Robertson, "High dielectric constant gate oxides for metal oxide Si transistors," *Reports on Progress in Physics*, vol. 69, no. 2, pp. 327–396, 2006.

[22] G. Bersuker, J. Sim, C. S. Park, C. D. Young, S. Nadkarni, R. Choi, and B.-H. Lee, "Mechanism of Electron Trapping and Characteristics of Traps in HfO_2 Gate Stacks," *IEEE Transactions on Device and Materials Reliability*, vol. 7, no. 1, pp. 138–145, 2007.

[23] G. Ribes, J. Mitard, M. Denais, S. Bruyere, F. Monsieur, C. Parthasarathy, E. Vincent, and G. Ghibaudo, "Review on high-κ dielectrics reliability issues," *IEEE Transactions on Device and Materials Reliability*, vol. 5, no. 1, pp. 5–19, 2005.

[24] S. Zafar, A. Kumar, E. Gusev, and E. Cartier, "Threshold voltage instabilities in high-κ gate dielectric stacks," *IEEE Transactions on Device and Materials Reliability*, vol. 5, no. 1, pp. 45–64, 2005.

[25] R. Degraeve, M. Aoulaiche, B. Kaczer, P. Roussel, T. Kauerauf, S. Sahhaf, and G. Groeseneken, "Review of reliability issues in high-κ/metal gate stacks," in *IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, 2008, pp. 1–6.

[26] X. Garros, L. Brunet, M. Rafik, J. Coignus, G. Reimbold, E. Vincent, A. Bravaix, and F. Boulanger, "PBTI mechanisms in La containing Hf-based oxides assessed by very Fast IV measurements," in *Proc. IEEE International Electron Devices Meeting (IEDM)*, 2010, pp. 4.6.1–4.6.4.

[27] E. Cartier and A. Kerber, "Stress-induced leakage current and defect generation in nFETs with HfO_2/TiN gate stacks during positive-bias temperature stress," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, 2009, pp. 486–492.

[28] D. Ioannou, S. Mitt, and G. LaRosa, "Positive Bias Temperature Instability Effects in advanced High-κ/Metal Gate NMOSFETs," in *Proc. IEEE IEEE International Integrated Reliability Workshop Final Report (IIRW)*, 2008, pp. 55–57.

[29] S. Krishnan, E. Cartier, J. Stathis, M. Chudzik, and A. Kerber, "Mechanistic understanding of Breakdown and Bias Temperature Instability in High-κ Metal devices using inline Fast Ramped Bias Test," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, 2011, pp. 4A.5.1–4A.5.5.

[30] S. Oktyabrsky and P. Ye, *Fundamentals of III-V Semiconductor MOSFETs*, 1st ed. Springer Publishing Company, Incorporated, 2010.

[31] T. Nigam, A. Kerber, and P. Peumans, "Accurate model for time-dependent dielectric breakdown of high-κ metal gate stacks," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, 2009, pp. 523–530.

[32] E. Cartier, B. Linder, V. Narayanan, and V. Paruchuri, "Fundamental understanding and optimization of PBTI in nFETs with SiO_2/HfO_2 gate stack," in *Proc. IEEE International Electron Devices Meeting (IEDM)*, 2006, pp. 1–4.

[33] A. Kerber, K. Maitra, A. Majumdar, M. Hargrove, R. Carter, and E. Cartier, "Characterization of Fast Relaxation During BTI Stress in Conventional and Advanced CMOS Devices With Gate Stacks," *IEEE Transactions on Electron Devices*, vol. 55, no. 11, pp. 3175–3183, 2008.

[34] K. J. Laidler, *Chemical Kinetics*, 3rd ed. Prentice Hall Publishing Company, Incorporated, 1987.

[35] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th ed. McGraw-Hill Publishing Company, Incorporated, 2001.

[36] M. Masuduzzaman, A. Islam, and M. Alam, "Exploring the Capability of Multifrequency Charge Pumping in Resolving Location and Energy Levels of Traps Within Dielectric," *IEEE Transactions on Electron Devices*, vol. 55, no. 12, pp. 3421–3431, 2008.

[37] M. Houssa, M. Tuominen, M. Naili, V. Afanasev, A. Stesmans, S. Haukka, and M. Heyns, "Trap-assisted tunneling in high permittivity gate dielectric stacks," *Journal of Applied Physics*, vol. 87, no. 12, pp. 8615–8620, 2000.

[38] M. Houssa, A. Stesmans, and M. M. Heyns, "Model for the trap-assisted tunnelling current through very thin SiO_2/ZrO_2 gate dielectric stacks," *Semiconductor Science and Technology*, vol. 16, no. 6, pp. 427–432, 2001.

[39] H. Kukner, S. Khan, P. Weckx, P. Raghavan, S. Hamdioui, B. Kaczer, F. Catthoor, L. V. der Perre, R. Lauwereins, and G. Groeseneken, "Comparison of Reaction-Diffusion and Atomistic Trap-Based BTI Models for Logic Gates," *IEEE Transactions on Device and Materials Reliability*, vol. 14, no. 1, pp. 182–193, March 2014.

[40] K. Kang, H. Kufluoglu, K. Roy, and M. A. Alam, "Impact of Negative-Bias Temperature Instability in Nanoscale SRAM Array: Modeling and Analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 26, no. 10, pp. 1770–1781, Oct 2007.

[41] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, "Matching properties of MOS transistors," *IEEE Journal of Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, Oct 1989.

[42] T. Naphade, K. Roy, and S. Mahapatra, "A Novel Physics-Based Variable NBTI Simulation Framework from Small Area Devices to 6T-SRAM," in *Proc. IEEE International Electron Devices Meeting (IEDM)*, 2013, pp. 33.6.1–33.6.4.

[43] S. Rauch, "Review and Reexamination of Reliability Effects Related to NBTI-Induced Statistical Variations," *IEEE Transactions on Device and Materials Reliability*, vol. 7, no. 4, pp. 524–530, 2007.

[44] B. Kaczer, T. Grasser, P. Roussel, J. Franco, R. Degraeve, L. A. Ragnarsson, E. Simoen, G. Groeseneken, and H. Reisinger, "Origin of NBTI variability in deeply scaled pFETs," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, 2010, pp. 26–32.

[45] *HSPICE Circuit Simulator*. Mountain View, California: Synopsys, Inc. [Online]. Available: http://www.synopsys.com/

[46] *Predictive Technology Model*. [Online]. Available: http://ptm.asu.edu/

[47] T. Grasser, *Bias Temperature Instability for Devices and Circuits*, 1st ed. Springer Publishing Company, Incorporated, 2014.

[48] M. Yabuuchi, R. Kishida, and K. Kobayashi, "Correlation between BTI-induced degradations and process variations by measuring frequency of ROs," in *Proc. IEEE International Meeting for Future of Electron Devices, Kansai (IMFEDK)*, June 2014, pp. 1–2.

[49] M. K. Hassan, C. H. Ho, and K. Roy, "Stochastic Modeling of Positive Bias Temperature Instability in High- (kappa) Metal Gate nMOSFETs,," *IEEE Transactions on Electron Devices*, vol. 61, no. 7, pp. 2243–2249, 2014.

[50] *MATLAB and Statistics Toolbox Release 2013b*. The MathWorks, Inc., Natick, Massachusetts, 2013. [Online]. Available: https://www.mathworks.com/products/

[51] D. B. M. Klaassen, "A unified mobility model for device simulationI. Model equations and concentration dependence," *Solid-State Electronics*, vol. 35, no. 7, pp. 953 – 959, 1992.

[52] M. J. van Dort, P. H. Woerlee, and A. J. Walker, "A simple model for quantisation effects in heavily-doped silicon MOSFETs at inversion conditions," *Solid-State Electronics*, vol. 37, no. 3, pp. 411 – 414, 1994.

[53] M. Toledano-Luque and et al, "Degradation of time dependent variability due to interface state generation," in *Proc. Symposium on VLSI Technology*, 2013, p. T190T191.

[54] A. Asenov, "Random Dopant Induced Threshold Voltage Lowering and Fluctuations in Sub-0.1 μm MOSFET's: A 3-D Atomistic Simulation Study," *IEEE Transactions on Electron Devices*, vol. 45, no. 12, pp. 2505–2513, Dec 1998.

[55] G. Panagopoulos and K. Roy, "A physics-based 3-D analytical model for RDF induced threshold voltage variations,," *IEEE Transactions on Electron Devices*, vol. 58, no. 2, pp. 392–403, 2011.

[56] R. Degraeve, B. Govoreanu, B. Kaczer, J. V. Houdt, and G. Groeseneken, "Measurement and statistical analysis of single trap current-voltage characteristics in ultrathin SiON," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, April 2005, pp. 360–365.

[57] D. M. Ramo, J. L. Gavartin, A. L. Shluger, and G. Bersuker, "Spectroscopic properties of oxygen vacancies in monoclinic HfO_2 calculated with periodic and embedded cluster density functional theory," *Physical Review B*, vol. 75, p. 205336, May 2007. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevB.75.205336

[58] R. Degraeve, M. Cho, B. Govoreanu, B. Kaczer, M. B. Zahid, J. V. Houdt, M. Jurczak, and G. Groeseneken, "Trap Spectroscopy by Charge Injection and Sensing (TSCIS): A quantitative electrical technique for studying defects in dielectric stacks," in *Proc. IEEE International Electron Devices Meeting (IEDM)*, Dec 2008, pp. 1–4.

[59] G. Darbandy and et al., "Analytical Modeling of Direct Tunneling Current Through Gate Stacks for the Determination of Suitable High-κ Dielectrics for Nanoscale Double-Gate MOSFETs," *Semiconductor Science and Technology*, vol. 26, no. 4, pp. 1–8, 2011.

[60] J. Stathies, "Percolation models for gate oxide breakdown," *Journal of Applied Physics*, vol. 86, no. 10, pp. 5757–5766, 1995.

[61] J. Sune, S. Tous, and E. Wu, "Analytical Cell-Based Model for the Breakdown Statistics of Multilayer Insulator Stacks," *IEEE Electron Device Letters*, vol. 30, no. 12, pp. 1359–1361, 2009.

[62] G. Bersuker and et al, "Breakdown in the metal/high-κ gate stack: Identifying the weak link in the multilayer dielectric," in *Proc. IEEE International Electron Devices Meeting (IEDM)*, 2008, pp. 1–4.

[63] K. S. Yew, D. S. Ang, and G. Bersuker, "Bimodal Weibull Distribution of Metal/High- κ Gate Stack TDDB-Insights by Scanning Tunneling Microscopy," *IEEE Electron Device Letters*, vol. 33, no. 2, pp. 146–148, Feb 2012.

[64] L. Pantisano and K. P. Cheung, "Stress-induced leakage current (SILC) and oxide breakdown: are they from the same oxide traps?" *IEEE Transactions on Device and Materials Reliability*, vol. 1, no. 2, pp. 109–112, Jun 2001.

[65] D. DiMaria and E. Cartier, "Mechanism for Stress-Induced Leakage Currents in Thin Silicon Dioxide Films," *Journal of Applied Physics*, vol. 78, no. 6, pp. 3883–3894, 1995.

[66] F. Crupi, R. Degraeve, A. Kerber, D. H. Kwak, and G. Groeseneken, "Correlation between Stress-Induced Leakage Current (SILC) and the HfO_2 bulk trap density in a SiO_2/HfO_2 stack," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, April 2004, pp. 181–187.

[67] S. Pae and et al., "Characterization of SILC and its end-of-life reliability assessment on 45NM high-κ and metal-gate technology," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, April 2009, pp. 499–504.

[68] C. D. Young and et al., "New insights into SILC-based life time extraction," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, April 2012, pp. 5D.3.1–5D.3.5.

[69] M. Alam, K. Roy, and C. Augustine, "Reliability- and Process-variation aware design of integrated circuits A broader perspective," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, 2011, pp. 4A.1.1 – 4A.1.11.

[70] A. E. Islam, H. Kufluoglu, D. Varghese, S. Mahapatra, and M. A. Alam, "Recent Issues in Negative-Bias Temperature Instability: Initial Degradation, Field Dependence of Interface Trap Generation, Hole Trapping Effects, and Relaxation," *IEEE Transactions on Electron Devices*, vol. 54, no. 9, pp. 2143–2154, Sept 2007.

[71] K. Joshi, S. Mukhopadhyay, N. Goel, and S. Mahapatra, "A consistent physical framework for N and P BTI in HKMG MOSFETs," in *Proc. IEEE International Reliability Physics Symposium (IRPS)*, April 2012, pp. 5A.3.1–5A.3.10.

[72] M. A. Alam, "A critical examination of the mechanics of dynamic NBTI for PMOSFETs," in *Proc. IEEE International Electron Devices Meeting (IEDM)*, Dec 2003, pp. 14.4.1–14.4.4.

[73] K. Kang, S. P. Park, K. Roy, and M. A. Alam, "Estimation of statistical variation in temporal NBTI degradation and its impact on lifetime circuit performance," in *Proc. IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, Nov 2007, pp. 730–734.

[74] Y. Nakagome, M. Horiguchi, T. Kawahara, and K. Itoh, "Review and future prospects of low-voltage RAM circuits," *IBM Journal of Research and Development*, vol. 47, no. 5.6, pp. 525–552, Sept 2003.

[75] P. Pouyan, E. Amat, and A. Rubio, "Process variability-aware proactive reconfiguration technique for mitigating aging effects in nano scale SRAM lifetime," in *proc. IEEE VLSI Test Symposium (VTS)*, April 2012, pp. 240–245.

[76] A. Bansal and et al., "Impacts of NBTI and PBTI on SRAM static/dynamic noise margins and cell failure probability," *Microelectronics Reliability*, vol. 49, no. 6, pp. 642 – 649, 2009, iMAPS-CPMT Poland 2008.

[77] B. Cheng, A. R. Brown, and A. Asenov, "Impact of NBTI/PBTI on SRAM Stability Degradation," *IEEE Electron Device Letters*, vol. 32, no. 6, pp. 740–742, June 2011.

[78] M. Sharifkhani and M. Sachdev, "SRAM Cell Stability: A Dynamic Perspective," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 2, pp. 609–619, Feb 2009.

[79] C. H. Ho, M. K. Hassan, S. Y. Kim, and K. Roy, "Analysis of Stability Degradation of SRAMs Using a Physics-Based PBTI Model," *IEEE Electron Device Letters*, vol. 35, no. 9, pp. 951–953, Sept 2014.

[80] M. K. Hassan and K. Roy, "Investigation of Dependence between Time-zero and Time-dependent Variability in High-κ NMOS Transistors," *arXiv:1609.05947*. [Online]. Available: https://arxiv.org/abs/1609.05947

[81] D. Khalil, M. Khellah, N. S. Kim, Y. Ismail, T. Karnik, and V. K. De, "Accurate Estimation of SRAM Dynamic Stability," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 16, no. 12, pp. 1639–1647, Dec 2008.

[82] S. Mukhopadhyay, H. Mahmoodi, and K. Roy, "Modeling of failure probability and statistical design of SRAM array for yield enhancement in nanoscaled CMOS," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 12, pp. 1859–1880, Dec 2005.

[83] R. H. Yan, A. Ourmazd, and K. F. Lee, "Scaling the Si MOSFET: from bulk to SOI to bulk," *IEEE Transactions on Electron Devices*, vol. 39, no. 7, pp. 1704–1710, Jul 1992.

[84] J. T. Park and J. P. Colinge, "Multiple-gate SOI MOSFETs: device design guidelines," *IEEE Transactions on Electron Devices*, vol. 49, no. 12, pp. 2222–2229, Dec 2002.

[85] B. S. Doyle, S. Datta, M. Doczy, S. Hareland, B. Jin, J. Kavalieros, T. Linton, A. Murthy, R. Rios, and R. Chau, "High performance fully-depleted tri-gate CMOS transistors," *IEEE Electron Device Letters*, vol. 24, no. 4, pp. 263–265, April 2003.

[86] *International Technology Roadmap for Semiconductors*, 2013. [Online]. Available: http://www.itrs2.net/itrs-reports.html

[87] *Spectre Circuit Simulator*. San Jose, California: Cadence Design Systems, Inc. [Online]. Available: https://www.cadence.com/

[88] S. Natarajan and et al., "A 14nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588 μm^2 SRAM cell size," in *Proc. IEEE International Electron Devices Meeting (IEDM)*, Dec 2014, pp. 3.7.1–3.7.3.

[89] J. G. Fossum, J. W. Yang, and V. P. Trivedi, "Suppression of corner effects in triple-gate MOSFETs," *IEEE Electron Device Letters*, vol. 24, no. 12, pp. 745–747, Dec 2003.

[90] H. Kufluoglu and M. A. Alam, "Theory of interface-trap-induced NBTI degradation for reduced cross section MOSFETs," *IEEE Transactions on Electron Devices*, vol. 53, no. 5, pp. 1120–1130, May 2006.

[91] H. Lee, C. H. Lee, D. Park, and Y. K. Choi, "A study of negative-bias temperature instability of SOI and body-tied FinFETs," *IEEE Electron Device Letters*, vol. 26, no. 5, pp. 326–328, May 2005.

[92] C. D. Young, A. Neugroschel, K. Majumdar, K. Matthews, Z. Wang, and C. Hobbs, "Investigation of negative bias temperature instability dependence on fin width of silicon-on-insulator-fin-based field effect transistors," *Journal of Applied Physics*, vol. 117, no. 3, pp. 034 501–034 508, January 2015.

[93] S. K. Gupta, G. Panagopoulos, and K. Roy, "NBTI in n-Type SOI Access FinFETs in SRAMs and Its Impact on Cell Stability and Performance," *IEEE Transactions on Electron Devices*, vol. 59, no. 10, pp. 2603–2609, Oct 2012.

[94] P. C. Feijoo, T. Kauerauf, M. Toledano-Luque, M. Togo, E. S. Andres, and G. Groeseneken, "Time-Dependent Dielectric Breakdown on Subnanometer EOT nMOS FinFETs," *IEEE Transactions on Device and Materials Reliability*, vol. 12, no. 1, pp. 166–170, March 2012.

[95] B. Kaczer, R. Degraeve, M. Rasras, K. V. d. Mieroop, P. J. Roussel, and G. Groeseneken, "Impact of MOSFET gate oxide breakdown on digital circuit operation and reliability," *IEEE Transactions on Electron Devices*, vol. 49, no. 3, pp. 500–506, Mar 2002.

APPENDICES

A. NBTI TRAP GENERATION DUE TO 2D DIFFUSION

In case of tri-gate FinFET devices, the corners can introduce 2D diffusion of hydrogen. Under NBTI stress, every free hydrogen molecule is associated with one interface trap. Fig. A.1 shows the 2D diffusion profile of hydrogen and it widens as $\sqrt{D_{H_2}t}$ over time [90]. The traps generated due to 2D-H_2 diffusion at the corners can be given as:

$$\begin{aligned} N_{IT}^{2D}(t) &= \frac{1}{4}\int_0^{\sqrt{D_{H_2}}} N_{H_2}(0)(1-\frac{r}{\sqrt{D_{H_2}t}})2\pi \mathrm{d}r \\ &= \frac{\pi}{2}\int_0^{\sqrt{D_{H_2}}} N_{H_2}(0)(1-\frac{r}{\sqrt{D_{H_2}t}})\mathrm{d}r \\ &= \frac{\pi}{2}N_{H_2}(0)([\frac{r^2}{2}]_0^{\sqrt{D_{H_2}}} - \frac{1}{\sqrt{D_{H_2}}}[\frac{r^3}{3}]_0^{\sqrt{D_{H_2}}}) \\ &= \frac{\pi}{2}N_{H_2}(0)(\frac{D_{H_2}}{2}(t) - \frac{D_{H_2}}{3}(t)) \\ &= \frac{\pi}{12}N_{H_2}(0)D_{H_2}t \end{aligned} \tag{A.1}$$

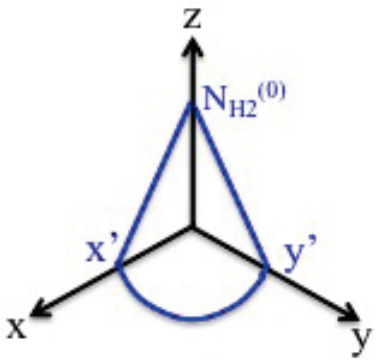


Fig. A.1. 2D diffusion profile of H_2 due to NBTI. $x' = y' = \sqrt{DH_2t}$ [90].

where, the parameters in (A.1) have their usual meanings. In this analysis, N_0, $N_H^{(0)}$, and $N_{H_2}^{(0)}$ are the Si-H bond density available before the NBTI stress is applied, H and H_2 density available for diffusion at the Si/Dielectric interface, respectively. From the analysis of 1D R-D model presented in [5], we can write the following expression for 2D diffusion:

$$\frac{k_f}{k_r} N_0 = N_{IT}^{2D} N_H(0) \tag{A.2}$$

$$k_1 = \frac{N_H^2(0)}{N_{H_2}(0)} \tag{A.3}$$

where, k_1 is a constant. Combining (A.2) and (A.3), we get:

$$N_{H_2}(0) = k_2 (\frac{k_f}{k_r} N_0)^2 (\frac{1}{N_{IT}^{2D}})^2 \tag{A.4}$$

where, k_2 is also a constant. Plugging in $N_{H_2}(0)$ from (A.4) in to (A.1), we determine the total number of interface traps generated due to 2D diffusion:

$$N_{IT}^{2D}(t) = \frac{\pi}{12} (\frac{k_f}{k_r} N_0)^2 (\frac{1}{N_{IT}^{2D}}(t))^2 (\frac{1}{k_2} D_{H_2} t) \tag{A.5}$$

Since, $\frac{1}{k_2}$ is a constant, (A.5) can be re-written as:

$$N_{IT}^{2D}(t) \propto (\frac{\pi}{12})^{\frac{1}{3}} (\frac{k_f}{k_0} N_0)^{\frac{2}{3}} (D_{H_2})^{\frac{1}{3}} t^{\frac{1}{3}} \tag{A.6}$$

Since the NBTI power exponent due to H_2 diffusion is $\frac{1}{6}$ [5], we observe from the above equation that $N_{IT}^{2D}(t) \propto t^{2n}$. Therefore, the $Si - H$ bonds at the corners are expected to degrade at a higher rate compared to the other parts of the Si/Dielectric interface.

B. SAMPLE CODES

B.1 nFinFET with Rectangular/Rounded Corners Fin: TCAD Sentaurus Structure Editor Code

```
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;          3D nFinFET (Rectangular Fin, Rounded Corners)          ;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
; Reinitializing DEVISE
(sde:clear)
; Selecting default Boolean expression
(sdegeo:set-default-boolean "BAB")

;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;                  DEFINITIONS - PARAMETERS                      ;;;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
(define DopSub "BoronActiveConcentration");
(define DopSD  "ArsenicActiveConcentration");
;   - Substrate, Source/Drain doping levels
(define SubDop  1e17 )    ; [1/cm3]
(define SDDop   5e20 )    ; [1/cm3]

;-----------------------------------------------------------------------
; Setting parameters
; - lateral
(define Lg    14.4e-3)              ; [um] Gate length
```

```
(define Wg    8e-3)                ; [um] Gate width
(define Lsp   3.5e-3)              ; [um] Spacer length
(define Lreox 1.5e-3)              ; [um] Reox length
(define SD_delta 2e-3)             ; [um] Tune S/D contact locations
(define Treox 2e-3)                ; [um] Reox2 thickness
(define Ltot 60e-3 )               ; [um] Lateral extend (each side)
(define del_Y 6e-3)                ; [um] Additional Y-extention


;--------------------------------------------------------------------
; - layers
(define Tsub 25e-3)          ; [um] Substrate thickness
(define Tsox  0.4e-3)        ; [um] Gate Si oxide thickness
(define Thfox  2e-3)         ; [um] Gate Hf oxide thickness
(define Tgate 30e-3)         ; [um] Metal gate thickness
(define Tfin 8e-3)       ; [um] Fin Thickness
(define Tfh 84e-3)       ; [um] 2xFin height
(define Tbody 8e-3)       ; [um] Body Thickness
(define Lsub 116e-3)       ; [um] Substrate Length
(define Wsub 116e-3)      ; [um] Substrate Width (for 70nm gate pitch)

;--------------------------------------------------------------------
;   - spacer rounding
(define fillet-radius 0.001)   ; [um] Rounding radius

;   - pn junction resolution
(define XjExt  0.005)          ; [um] Extension depth
(define XjSD   0.042)          ; [um] SD Junction depth
```

```
; Derived quantities
(define Xg (/ Lg 2))
(define Yg (/ Wsub 2))
(define Tfg (/ Tfin 2))
(define Tfhg (/ Tfh 2))        ;Fin Height
(define Hg (+ Tfhg 5e-3))      ; Gate Height
(define Tox (+ Tsox Thfox))

;-----------------------------------------------------------------
; Creating regions
(sdegeo:create-cuboid (position  (* Ltot -1) Yg (* Tsub -1))
(position Ltot (* Yg -1) 0.0) "Silicon" "Si_Substrate")

(sdegeo:create-cuboid (position  (* Tfg -1) Yg 0.0)
(position Tfg (* Yg -1) Tfh) "StrainedSilicon" "Fin_1")
(sdegeo:fillet (list (car (find-vertex-id (position  (* Tfg -1)
Yg Tfh))) (car (find-vertex-id (position  (* Tfg -1) (* Yg -1) Tfh)))
(car (find-vertex-id (position  Tfg Yg Tfh)))
(car (find-vertex-id (position Tfg (* Yg -1) Tfh))))
(* fillet-radius 0.825))
(sdegeo:create-cuboid (position  (* Ltot -1) Yg 0.0)
(position Ltot (* Yg -1) Tfhg) "SiO2" "Native_Oxide")
(sdegeo:create-cuboid (position  (* (+ Tfg Tsox) -1) (+ Xg 0) Tfhg)
(position (+ Tfg Tsox) (* (+ Xg 0) -1) (+ Tfh Tsox)) "SiO2" "Sox")
(sdegeo:fillet (list (car (find-edge-id (position  (* (+ Tfg Tsox) -1)
0 (+ Tfh Tsox)))) (car (find-edge-id (position (+ Tfg Tsox) 0
(+ Tfh Tsox))))) (* fillet-radius 1.1))
(sdegeo:create-cuboid (position  (* (+ Tfg Tox) -1) (+ Xg 0) Tfhg)
(position (+ Tfg Tox) (* (+ Xg 0) -1) (+ Tfh Tox)) "HfO2" "Hfox")
```

```
(sdegeo:fillet (list (car (find-edge-id (position  (* (+ Tfg Tox) -1)
0 (+ Tfh Tox)))) (car (find-edge-id (position (+ Tfg Tox) 0
(+ Tfh Tox))))) (* fillet-radius 2))
(sdegeo:create-cuboid (position  (* (/ Ltot 1) -1) Xg Tfhg)
(position (/ Ltot 1) (* Xg -1) (+ Tfhg Hg)) "TiN" "Metal_Gate")

;-----------------------------------------------------------------
; Creating Re-oxidation and Spacer regions
(sdegeo:create-cuboid (position  (* (/ Ltot 1) -1) (+ Xg Lreox)
Tfhg) (position (/ Ltot 1) (* (+ Xg Lreox) -1) (+ Tfhg Hg)) "SiO2"
"Gate_Reox1")

(sdegeo:create-cuboid (position  (* (/ Ltot 1) -1) (+ (+
Xg Lreox) Lsp) Tfhg) (position (/ Ltot 1) (* (+ (+ Xg Lreox)
Lsp) -1) (+ Tfhg Hg)) "Si3N4" "N_Spacer")

;-----------------------------------------------------------------
; - rounding spacer
(sdegeo:fillet (list (car (find-vertex-id (position  (/ Ltot 1) (+ (+
Xg Lreox) Lsp) (+ Tfhg Hg)))) (car (find-vertex-id (position
(/ Ltot 1) (* (+ (+ Xg Lreox) Lsp) -1) (+ Tfhg Hg)))) (car
(find-vertex-id (position  (* (/ Ltot 1) -1) (+ (+ Xg Lreox) Lsp)
(+ Tfhg Hg)))) (car (find-vertex-id (position (* (/ Ltot 1) -1)
(* (+ (+ Xg Lreox) Lsp) -1) (+ Tfhg Hg))))) fillet-radius)

;-----------------------------------------------------------------
; Define Contacts
(sdegeo:define-contact-set "gate"      4.0
(color:rgb 1.0 0.0 0.0 ) "##" )
```

```
(sdegeo:define-contact-set "drain"     4.0
(color:rgb 1.0 0.0 0.0 ) "||" )
(sdegeo:define-contact-set "source"    4.0
(color:rgb 1.0 0.0 0.0 ) "==" )
(sdegeo:define-contact-set "substrate" 4.0
(color:rgb 1.0 0.0 0.0 ) "<><>" )

;-----------------------------------------------------------------
;------------------------ Set Contacts-------------------------
; Substrate and Gate
(sdegeo:set-current-contact-set "substrate")
(sdegeo:define-3d-contact (list (car (find-face-id
(position 0 0 (* Tsub -1))))) "substrate")
(sdegeo:set-current-contact-set "gate")
(sdegeo:define-3d-contact (list (car (find-face-id
(position 0 0 (+ Tfhg Hg))))) "gate")


; Drain and Source
(sdegeo:create-cuboid (position  (* Tfg -1) Yg Tfh)
(position Tfg (* Yg -1) (+ Tfh 5e-3)) "Metal" "DS")
(sdegeo:delete-region (list (car (find-body-id
(position 0 Yg (+ Tfh 2e-3))))))

(sdegeo:set-current-contact-set "drain")
(sdegeo:define-3d-contact (list (car (find-face-id
(position 0 (* (+ (+ (+ Xg Lreox) Lsp) 5e-3) -1) Tfh)))) "drain")
(sdegeo:set-current-contact-set "source")
(sdegeo:define-3d-contact (list (car (find-face-id
```

```
(position 0 (+ (+ (+ Xg Lreox) Lsp) 5e-3) Tfh)))) "source")

;-----------------------------------------------------------------
;--------------------- Set Doping Profiles---------------------

; Constant Profile (Bulk and Fins)

(sdedr:define-constant-profile "Const.Bulk" DopSub SubDop)
(sdedr:define-constant-profile-region "PlaceCD.Bulk" "Const.Bulk"
"Si_Substrate")
(sdedr:define-constant-profile-region "PlaceCD.Fin1" "Const.Bulk"
"Fin_1")

; Analytic Profile (Drain and Source)
(sdedr:define-refinement-window "BaseLine.Drain" "Rectangle"
(position (* (+ Tfg 2e-3) -1) (* (+ Xg (+ Lreox (+ SD_delta Lsp))) -1)
Tfh)  (position (+ Tfg 2e-3) (* (+ Yg 2e-3) -1)  Tfh))
(sdedr:define-refinement-window "BaseLine.Source" "Rectangle"
(position (* (+ Tfg 2e-3) -1) (* (+ Xg (+ Lreox (+ SD_delta Lsp))) 1)
Tfh)  (position (+ Tfg 2e-3) (* (+ Yg 2e-3) 1)  Tfh))
(sdedr:define-gaussian-profile "Gauss.SourceDrain" DopSD
"PeakPos" 0.0 "PeakVal" SDDop "ValueAtDepth" SubDop "Depth"
XjSD "Gauss" "Factor" 0.1)
(sdedr:define-analytical-profile-placement "PlaceAP.Drain"
"Gauss.SourceDrain" "BaseLine.Drain" "Negative" "NoReplace" "Eval")
(sdedr:define-analytical-profile-placement "PlaceAP.Source"
"Gauss.SourceDrain" "BaseLine.Source" "Negative" "NoReplace" "Eval")

;-----------------------------------------------------------------
```

```
;--------------------- Defining Meshing Strategy ----------------------

; Global Mesh
(sedr:define-refinement-window "RefWin.Global" "Cuboid"
(position  (* Ltot -1) Yg (* Tsub -1)) (position Ltot (* Yg -1)
(+ Tfhg Hg)))
(sedr:define-refinement-size "RefDef.Global" 0.3
0.3  0.3 0.1 0.1 0.1 )
(sedr:define-refinement-function "RefDef.Global"
"DopingConcentration" "MaxTransDiff" 0.1)
(sedr:define-refinement-placement "Place.Global"
"RefDef.Global" "RefWin.Global" )

; SiO2
(sedr:define-refinement-size "RefDef.Sox" 0.002 0.002
0.001 0.001 0.001 0.0004 )
(sedr:define-refinement-function "RefDef.Sox"
"DopingConcentration" "MaxTransDiff" 0.1)
(sedr:define-refinement-region "Place.Sox" "RefDef.Sox" "Sox" )

; HfO2
(sedr:define-refinement-size "RefDef.Hfox" 0.002 0.002
0.002 0.001 0.001 0.001 )
(sedr:define-refinement-function "RefDef.Hfox"
"DopingConcentration" "MaxTransDiff" 0.1)
(sedr:define-refinement-region "Place.Hfox" "RefDef.Hfox" "Hfox" )

;Channel Region
(sedr:define-refinement-size "RefDef.Channel" 0.006 0.006 0.006
```

```
0.001 0.001 0.001 )
(sdedr:define-refinement-function "RefDef.Channel"
"DopingConcentration" "MaxTransDiff" 0.1)
(sdedr:define-refinement-region "Place.Channel"
"RefDef.Channel" "Fin_1" )

;Top Region
(sdedr:define-refinement-window "RefWin.Top" "Cuboid"
(position  (* (/ Ltot 3) -1) (+ (+ Xg Lreox) Lsp) Tfhg)
(position (/ Ltot 3) (* (+ (+ Xg Lreox) Lsp) -1) (+ Tfhg Hg)))
(sdedr:define-refinement-size "RefDef.Top" 0.003  0.003
0.003 0.001 0.001 0.001 )
(sdedr:define-refinement-function "RefDef.Top"
"DopingConcentration" "MaxTransDiff" 0.1)
(sdedr:define-refinement-placement "Place.Top"
"RefDef.Top" "RefWin.Top" )

;-----------------------------------------------------------------
(sde:save-model "./NFin_3D")
(sde:build-mesh "snmesh" "" "./NFin_3D")
(sdeio:save-dfise-bnd (get-body-list) "./NFin_3D_msh.bnd")
;-----------------------------------------------------------------
```

B.2 Determine Percolation Paths: Matlab Code

```
% 3D Percolation Model for TDDB Analysis
%'Linear Search' to find nearest neighbors sharing at least one
%common edge
% This code can capture a 4-defects percolation paths,
% but can be modified to a different number of defect-paths
% Most of the recent papers suggest that SBD defect paths
% are mostly comprised of 3-4 defects in scaled techonologies
% tBD: time to breakdown:: NBD:critical breakdown defect density::
% XB, YB, and ZB: locations of the defects :: Tin: Matrix containing
% all the possible defect locations
%%
function [c, tBD, NBD, XB, YB, ZB, BD_M]=...
breakdown_path(BD_M, ct, Tin, c, tBD, NBD, XB, YB, ZB)

    [Xi, Yi, Zi] = ind2sub(size(Tin),find(Tin));

    tf1=isempty(find(Zi==1));
    tf2=isempty(find(Zi==2));
    tf3=isempty(find(Zi==3));
    tf4=isempty(find(Zi==4));

% check if defects are formed in all the layers
    TF=tf1|tf2|tf3|tf4;

    if(TF==1)
        return;
    else
```

```
% Extract the locations of the defects
        [Xi, Yi, Zi] = ind2sub(size(Tin),find(Tin));
% Form a vector using the locations
          D_M=[Xi'
               Yi'
               Zi'];

    for k=1:N_layers
        count_L(k)=sum(D_M(3,:)==k);
    end
    X=D_M(1,:);    Y=D_M(2,:);     Z=D_M(3,:);

% Searches for a continuous path and saves the data
for m=1:count_L(1)
 for n=count_L(1)+1:count_L(1)+count_L(2)
  if((abs(X(m)-X(n))<=1)&&(abs(Y(m)-Y(n))<=1))
   for p=count_L(1)+count_L(2)+1:count_L(1)+count_L(2)+count_L(3)
    if((abs(X(n)-X(p))<=1)&&(abs(Y(n)-Y(p))<=1))
     for r=count_L(1)+count_L(2)+count_L(3)+1:count_L(1)+...
count_L(2)+count_L(3)+count_L(4)
      if(((BD_M(X(m),Y(m),Z(m))==0)||(BD_M(X(n),Y(n),Z(n))==0)...
||(BD_M(X(p),Y(p),Z(p))==0)||(BD_M(X(r),Y(r),Z(r))==0))...
&&((abs(X(p)-X(r))<=1)&&(abs(Y(p)-Y(r))<=1)))
               tBD(c)=ct;
               NBD(c)=nnz(Tin);
               XB(c,:)=[X(m), X(n), X(p), X(r)];
       YB(c,:)=[Y(m), Y(n), Y(p), Y(r)];
       ZB(c,:)=[Z(m), Z(n), Z(p), Z(r)];
               BD_M(X(m),Y(m),Z(m))=ct;
```

```
      BD_M(X(n),Y(n),Z(n))=ct;
             BD_M(X(p),Y(p),Z(p))=ct;
      BD_M(X(r),Y(r),Z(r))=ct;
             c=c+1;
end; end; end; end; end; end; end; end;
```

VITA

VITA

Md. Khaled Hassan received the B.S. degree in Electrical & Electronic Engineering from Bangladesh University of Engineering & Technology in 2006 and M.S. in Electrical & Computer Engineering from the University of Illinois at Chicago in 2012. He is currently working toward the Ph.D. degree in Electrical Engineering at Purdue University, West Lafayette, IN.